Search for continuous gravitational waves: improving robustness versus instrumental artifacts

David Keitel* and Reinhard Prix†

Albert-Einstein-Institut, Callinstrasse 38, 30167 Hannover, Germany

Maria Alessandra Papa and Paola Leaci

Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Golm and
Callinstrasse 38, 30167 Hannover, Germany

Maham Siddiqi

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

The standard multi-detector $\mathcal{F}$-statistic for continuous gravitational waves is susceptible to false alarms from instrumental artifacts, for example monochromatic sinusoidal disturbances ("lines"). This vulnerability to line artifacts arises because the $\mathcal{F}$-statistic compares the signal hypothesis to a Gaussian-noise hypothesis, and hence is triggered by anything that resembles the signal hypothesis *more* than Gaussian noise. Various *ad-hoc* veto methods to deal with such line artifacts have been proposed and used in the past. Here we develop a Bayesian framework that includes an explicit alternative hypothesis to model disturbed data. We introduce a simple line model that defines lines as signal candidates appearing only in one detector. This allows us to explicitly compute the odds between the signal hypothesis and an extended noise hypothesis, resulting in a new detection statistic that is more robust to instrumental artifacts. We present and discuss results from Monte-Carlo tests on both simulated data and on detector data from the fifth LIGO science run. We find that the line-robust detection statistic retains the detection power of the standard $\mathcal{F}$-statistic in Gaussian noise, while it can be substantially more sensitive in the presence of line artifacts. This new statistic also equals or surpasses the performance of the popular $\mathcal{F}$-statistic consistency veto.

PACS numbers: 04.30.Tv, 04.80.Nn, 95.55.Ym, 97.60.Jd

I. INTRODUCTION

Spinning neutron stars with non-axisymmetric deformations are expected to emit quasi-monochromatic and long-lasting gravitational waves (GWs), commonly referred to as *continuous waves* (CWs). One of the main search methods for CWs was developed for ground-based detectors (such as LIGO [1], Virgo [2], GEO 600 [3]) and it is the so-called $\mathcal{F}$-statistic. The $\mathcal{F}$-statistic was originally derived as a maximum-likelihood detection statistic [4, 5]; it was later shown that it can also be derived as a Bayes factor using somewhat unphysical priors for the signal amplitude parameters [6].

In the context of CW searches, the GW data is reasonably well described by an underlying Gaussian noise distribution with additional non-Gaussian disturbances. The $\mathcal{F}$-statistic corresponds to a binary hypothesis test between a signal hypothesis and a Gaussian-noise hypothesis. As a consequence, it is possible to obtain large $\mathcal{F}$-statistic values due to non-Gaussian disturbances in the data, even if they are not well matched to the signal model. Large $\mathcal{F}$-statistic values only imply that the signal hypothesis is a *better* fit to the data than pure Gaussian noise, but they do not imply a *good* fit.

The most problematic instrumental artifacts for any specific analysis of GW data are typically those that resemble the signal family it searches for, i.e., disturbances with non-negligible projection onto the signal templates of a search. For example, searches for short transient signals, such as bursts (e.g., from core-collapse supernovae) or compact binary coalescences (CBCs), are most affected by "glitches" in the data, i.e., short broad-band disturbances [7–10].

On the other hand, searches for CW signals are mainly affected by so-called "lines", i.e., narrow-band disturbances that are present for a sizable fraction of the observation time. Examples include the so-called *mains* lines (i.e., lines at multiples of the 60 Hz electrical power system frequency), the resonance frequencies of the detector suspensions, and lines from digital components – see Ref. [11] for more details and a list of known instrumental lines identified in the data from the fifth LIGO science run (S5), and Refs. [8, 12] for line identification in Virgo data.

In this article, we apply a Bayesian model-selection approach using an additional alternative noise hypothesis for lines. Since the characteristics of the population of instrumental lines affecting CW searches are not well understood, we use a very simple line model. The model is based on an observed distinguishing feature of many lines, namely that they do not affect all detectors in the same way. Hence, we define our line model as any feature in the data that resembles the signal template *in*

---

* David.Keitel@aei.mpg.de
† Reinhard.Prix@aei.mpg.de

only one detector.

This approach can also be seen as adding a coincidence criterion to the coherent multi-detector $\mathcal{F}$-statistic. Such a method is applicable only to multi-detector CW searches, and in practice the most recent $\mathcal{F}$-statistic-based searches have all used data from multiple detectors. By employing this approach, we obtain a new line-robust detection statistic, generalizing the $\mathcal{F}$-statistic.

The plan of this paper is as follows: In Sec. II we give a short review of existing methods to deal with glitches and lines in GW data. In Sec. III we describe signal-noise hypotheses relevant for the detection problem at hand: the standard Gaussian-noise hypothesis, in Sec. III A, the CW signal hypothesis, in Sec. III B and a new simple line hypothesis in Sec. III C. In Sec. IV we use these hypotheses to construct two new detection statistics, a "line-veto" statistic and a more general line-robust statistic. We generalize the hypotheses and statistics to the case of semi-coherent searches in Sec. V. Next we discuss the choice of prior parameters for the line-robust statistic and we present a simple method, albeit somewhat *ad hoc*, to choose decent priors in Sec. VI. This concludes the analytical part of the paper. In Sec. VII we assess the performance of the new statistics through a series of numeric tests: on fully synthetic data in Sec. VII A and on LIGO S5 data in Sec. VII B. We summarize our findings in Sec. VIII and give a short outlook on applications and future generalizations of this approach. The appendix A contains a short derivation of the expected $\mathcal{F}$-statistic value under the simple line hypothesis.

II. EXISTING METHODS TO MITIGATE DETECTOR ARTIFACTS

The problem of non-Gaussian artifacts in the data affects both searches for long-lived signals (e.g., the CWs which are the topic of this paper) as well as searches for short-lived signals. Short-lived signals are expected from the late phase of the inspiral of binaries of compact objects such as neutron stars and black holes as well as from catastrophic events such as supernovae. For short-lived signal searches the artifacts that are responsible for an increase in the false alarm rates with respect to purely Gaussian noise manifest themselves as loud glitches in the time-domain data. On the other hand, for long-lived signal searches the most troublesome artifacts are broadly speaking those that appear in the Fourier spectra on the typical timescales of the search.

An interesting distinction in search pipelines is the order in which multi-detector coherence and coincidence are used. If the first step in the search is a coherent multi-detector statistic (as for the CW searches that we consider here), then the noise-artifact-mitigation strategy may use subsequent consistency checks on the statistics from the individual detectors. If, on the other hand, the first step was a single-detector search followed by a selection of triggers where only coincidence in the various detectors is required, an additional multi-detector coherent statistic can serve as an artifact-mitigation technique.

A wide range of methods have been developed in order to deal with instrumental artifacts. In the following subsections, we give a short review of such methods for CBC, burst and CW searches.

Generally, we can distinguish between two fundamentally different approaches to artifact mitigation: Bayesian model-selection and heuristic methods. The former is based on *explicit* alternative models whereas the latter consists in constructing *ad-hoc* statistics to detect certain observed deviations from the GW signal model, which in the Bayesian picture corresponds to a test against *implicit* (and often unknown) alternative hypotheses. A third "hybrid" approach uses Bayesian inference to directly construct empirical noise and signal likelihoods using the actual data and injections [13].

A. Instrumental glitches in burst and CBC searches

In searches for short-lived GW signals popular *ad-hoc* glitch-veto methods are the χ^2-veto [14], the null-stream veto [15], and signal amplitude consistency vetoes [16]. These (among others) are commonly used in CBC searches (e.g., see Refs. [17–19]) and in burst searches (e.g., see Refs. [20–22]).

For instance, in low-mass CBC searches the first step is a separate search in each detector. After a cut on single-detector χ^2 values, the glitch mitigation strategy consists in applying a coincidence criterion and then constructing on the surviving candidates a new multi-detector statistic. This folds in the single-detector statistics and χ^2 values. Significance thresholds are set based on Monte-Carlo studies on actual data and injections.

In searches for signals for which we lack a waveform model (i.e., generic bursts), a main multi-detector statistic is constructed that accounts appropriately for time delays and antenna responses of the different detectors to the same putative GW. This statistic is then augmented by other statistics (see Ref. [20] for details) specifically designed to further check for signal consistency across the detectors by means of appropriate veto conditions.

Various explicit glitch models have been considered, including Sine-Gaussians [23, 24] and wavelets [25], and Bayesian approaches have also been proposed to use these in constructing glitch-robust searches [23, 25]. Notably, Veitch & Vecchio [26] have defined a glitch model describing coincident single-detector candidates with independent amplitude parameters in different detectors. On the other hand, the signal model requires candidates to be both coincident and coherent across all detectors. Both hypotheses would fit a true signal equally well, but the glitch hypothesis would be weighed down by its larger prior volume ("Occam's razor"). In the case of glitches, however, the glitch hypothesis will generally provide a much better fit, allowing it to overcome its larger prior

volume.

B. Instrumental lines in CW searches

The most commonly used approaches to deal with instrumental lines in CW searches are all heuristic and can be summarized as follows:

Line cleaning: This is a widely used approach with many variants. It consists in effectively excluding frequency bands from the search when they are known or believed to be affected by instrumental lines. This could be either as a result of previous detector characterization work or because the frequency-domain data was flagged as particularly disturbed (referred to as *line flagging*). Among the examples of this approach are the LIGO/Virgo searches [11, 27–30].

A downside of this method is the relatively large fraction of the total frequency band it typically vetoes. For example, in Ref. [27] it vetoed a total of 270 Hz out of the 1140 Hz searched, i.e., $\sim 24\%$ of the data. Furthermore, this method is either limited to known instrumental lines or, when the line-flagging variant is used, its efficacy is limited to strong disturbances. Weaker disturbances can only be identified with time baselines much longer than the ones typically used by the line-flagging algorithms. Furthermore, the Fourier-transform-based line-flagging algorithm is not optimally suited to detect lines with non-constant frequency.

S-veto: This is a method to remove candidates from a (frequency and spindown dependent) region of the sky. This region is typically around the poles, where the corresponding signal templates are not well distinguished from typical instrumental line artifacts. This method was initially developed in PowerFlux [29] and subsequently adapted to $\mathcal{F}$-statistic searches [31].

The fraction of the total parameter space vetoed *a-priori* through this approach can again be quite large, for example about $\sim 30\%$ in Ref. [31].

$\mathcal{F}$-statistic consistency veto: If a candidate from a multi-detector search has a single-detector $\mathcal{F}$-statistic value exceeding its multi-detector $\mathcal{F}$-statistic, then it is vetoed as a likely instrumental line. This approach was described in more detail and tested in Ref. [11] and Refs. [28, 32]. This veto method will be recovered as a special case of the line-veto statistic developed in Sec. IV A.

The approach proposed in the present work is not a heuristic method as the ones described above. Instead, it shares some similarities to the glitch-robust method proposed in Ref. [26], but it differs in the following: In the incoherent CBC pipeline, any candidate is already required to be *coincident* between detectors, so the method of Ref. [26] adds the requirement of multi-detector *coherence* to distinguish GW signals from glitches. In our case, instead, we start from the *coherent* multi-detector $\mathcal{F}$-statistic and add a *coincidence* requirement to distinguish CW signals from (non-coincident) lines.

Currently we do not include coincident lines in the alternative hypothesis, as we expect that this would substantially weaken the detection power of this method. More work is required to deal with coincident lines, which trigger the same templates in multiple detectors. The approach taken here is that in a full CW search pipeline the non-coincident line model would serve as a cheap and simple "first line of defense" to reduce the number of spurious candidates, while more sophisticated steps can be applied to the surviving candidates in later steps.

III. HYPOTHESES ABOUT THE OBSERVED DATA

Let $x^X(t)$ be the time series of GW strain measured in a detector X, where we use $X, Y, \ldots$ as detector indices. Following the multi-detector notation from Refs. [5, 33], using boldface indicates a multi-detector vector, i.e., we write $\boldsymbol{x}(t)$ for the multi-detector data vector with components $x^X(t)$.

We will consider three different hypotheses about the observed data $\boldsymbol{x}$ and derive their posterior probabilities: the Gaussian noise hypothesis $\mathcal{H}_{\mathrm{G}}$, the CW signal hypothesis $\mathcal{H}_{\mathrm{S}}$ and a simple "line" hypothesis $\mathcal{H}_{\mathrm{L}}$.

A. The Gaussian noise hypothesis $\mathcal{H}_{\mathrm{G}}$

The Gaussian-noise hypothesis $\mathcal{H}_{\mathrm{G}}$ states that the measured multi-detector time series $\boldsymbol{x}(t)$ only contains stationary Gaussian noise, which we denote as $\boldsymbol{n}(t)$, i.e.,

$$\mathcal{H}_{\mathrm{G}} : \boldsymbol{x}(t) = \boldsymbol{n}(t) \,, \tag{1}$$

with a single-sided power-spectral density (PSD) $\boldsymbol{S}_n$ that is assumed to be known. The corresponding likelihood for measuring the data $\boldsymbol{x}$ can therefore be written as

$$P\left(\boldsymbol{x}|\mathcal{H}_{\mathrm{G}}\right) = \kappa \, e^{-\frac{1}{2}(\boldsymbol{x}|\boldsymbol{x})} \,, \tag{2}$$

where κ is a data-independent normalization constant, and the scalar product is defined as

$$(\boldsymbol{x}|\boldsymbol{y}) \equiv \sum_X \frac{1}{S_{\mathrm{n}}^X} \int_0^T x^X(t)\, y^X(t)\, dt \,, \tag{3}$$

assuming that the noise spectra S_{n}^X are uncorrelated between different detectors X and constant over the (narrow) frequency band of interest. For simplicity of notation we omit the sometimes customary notation of a

conditional "I" denoting all implicit and explicit model assumptions, i.e., we write $P(a|b)$ as a shortcut for $P(a|b, I)$, and $P(a)$ as an abbreviation for $P(a|I)$.

The posterior probability for $\mathcal{H}_\mathrm{G}$ given the observed data $\boldsymbol{x}$ follows from Bayes' theorem as

$$P\left(\mathcal{H}_\mathrm{G}|\boldsymbol{x}\right) = \frac{P\left(\mathcal{H}_\mathrm{G}\right)}{P\left(\boldsymbol{x}\right)}\,\kappa\,\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}|\boldsymbol{x})}\,, \tag{4}$$

where $P\left(\mathcal{H}_\mathrm{G}\right)$ is the prior probability for the Gaussian-noise hypothesis. The normalization $P\left(\boldsymbol{x}\right)$ depends on the full set of assumed hypotheses $\{\mathcal{H}_i\}$, i.e., $P\left(\boldsymbol{x}\right) = \sum_i P\left(\boldsymbol{x}|\mathcal{H}_i\right) P\left(\mathcal{H}_i\right)$, but in the following we will only consider the *odds* between different hypotheses, where this term drops out.

B. The CW signal hypothesis $\mathcal{H}_\mathrm{S}$

The hypothesis $\mathcal{H}_\mathrm{S}$ for CW signals [4, 34] states that the data $\boldsymbol{x}$ contains a CW signal $\boldsymbol{h}$ in addition to Gaussian noise $\boldsymbol{n}$, namely $\boldsymbol{x} = \boldsymbol{n} + \boldsymbol{h}$.

The signal model $\boldsymbol{h}$ depends on a number of (generally unknown) signal parameters. For practical reasons, we usually distinguish between the set of four *amplitude parameters* $\mathcal{A}$ and the remaining *phase-evolution parameters* λ, i.e., we write the CW signal family as $\boldsymbol{h}(t; \mathcal{A}, \lambda)$.

To fully specify the signal hypothesis, we therefore need a prior probability distribution $P\left(\mathcal{A}, \lambda|\mathcal{H}_\mathrm{S}\right)$ for the signal parameters, i.e.,

$$\begin{aligned}\mathcal{H}_\mathrm{S} : \boldsymbol{x}(t) = \boldsymbol{n}(t) &+ \boldsymbol{h}(t; \mathcal{A}, \lambda) \\ \text{with prior } &P\left(\mathcal{A}, \lambda|\mathcal{H}_\mathrm{S}\right)\,. \end{aligned} \tag{5}$$

The amplitude parameters $\mathcal{A}$ describe the signal amplitude h_0, the inclination angle ι, the polarization angle ψ and the initial phase ϕ_0. As first shown in Ref. [4], a particular parametrization $\mathcal{A}^\mu = \mathcal{A}^\mu(h_0, \cos\iota, \psi, \phi_0)$, with $\mu = 1\ldots 4$, allows one to write the signal model in the factorized form

$$\boldsymbol{h}(t; \mathcal{A}, \lambda) = \mathcal{A}^\mu\,\boldsymbol{h}_\mu(t; \lambda)\,, \tag{6}$$

in terms of four basis functions $\boldsymbol{h}_\mu(t; \lambda)$ and using the automatic summation convention over repeated indices.

In order to simplify the following discussion and notation, we follow the approach of Refs. [6, 35] and formally restrict ourselves to a single-template statistic in λ. This is equivalent to the assumption of known phase parameters, i.e., $\lambda = \lambda_\mathrm{s}$. This can be done without loss of generality, as for unknown $\lambda \in \mathbb{P}$ this analysis would apply for each template $\lambda_i \in \mathbb{P}$, and one would then marginalize over the prior parameter space $\mathbb{P}$. Studying this in further detail is outside the scope of the present work. We will therefore assume a prior of the form

$$P\left(\mathcal{A}, \lambda|\mathcal{H}_\mathrm{S}\right) = P\left(\mathcal{A}|\mathcal{H}_\mathrm{S}\right)\,\delta(\lambda - \lambda_\mathrm{s})\,, \tag{7}$$

and drop the phase-evolution parameters λ from the following expressions.

We can obtain the likelihood for *a particular* signal $\boldsymbol{h}(t; \mathcal{A})$ by noting that, according to $\mathcal{H}_\mathrm{S}$, the combination $[\boldsymbol{x} - \boldsymbol{h}(\mathcal{A})]$ is described by Gaussian noise. In fact, by inserting the signal factorization from Eq. (6) and by factoring out terms equivalent to the Gaussian noise likelihood from Eq. (4), we obtain

$$\begin{aligned}P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{S}, \mathcal{A}\right) &= \kappa\,\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{h}(\mathcal{A})|\boldsymbol{x}-\boldsymbol{h}(\mathcal{A}))} \\ &= \kappa\,\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}|\boldsymbol{x})}\,\mathrm{e}^{(\boldsymbol{x}|\mathcal{A}^\mu \boldsymbol{h}_\mu) - \frac{1}{2}(\mathcal{A}^\mu \boldsymbol{h}_\mu|\mathcal{A}^\nu \boldsymbol{h}_\nu)} \\ &= P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{G}\right) \exp\left[\mathcal{A}^\mu\,x_\mu - \frac{1}{2}\mathcal{A}^\mu \mathcal{M}_{\mu\nu}\mathcal{A}^\nu\right]\,, \end{aligned} \tag{8}$$

where we introduced the four projections x_μ of the data and the (symmetric positive-definite) matrix $\mathcal{M}_{\mu\nu}$ as

$$x_\mu \equiv (\boldsymbol{x}|\boldsymbol{h}_\mu) \quad \text{and} \quad \mathcal{M}_{\mu\nu} \equiv (\boldsymbol{h}_\mu|\boldsymbol{h}_\nu)\,. \tag{9}$$

The *marginal* likelihood $P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{S}\right)$ (sometimes referred to as "evidence") for the signal hypothesis from Eq. (5) can be obtained by marginalizing over the unknown amplitudes $\mathcal{A}$, namely

$$P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{S}\right) = \int P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{S}, \mathcal{A}\right) P\left(\mathcal{A}|\mathcal{H}_\mathrm{S}\right)\,d\mathcal{A}\,. \tag{10}$$

This integral can be solved analytically for certain choices of amplitude priors $P\left(\mathcal{A}|\mathcal{H}_\mathrm{S}\right)$. In particular, as discussed in Refs. [6, 35], for the (somewhat unphysical) prior that is uniform in $\mathcal{A}^\mu$, we can recover the standard $\mathcal{F}$-statistic, namely assuming

$$P\left(\{\mathcal{A}^\mu\}|\mathcal{H}_\mathrm{S}\right) = \begin{cases} C & \text{for} \quad h_0^4(\mathcal{A}) < \frac{70\,c_*}{\sqrt{|\mathcal{M}|}}\,. \\ 0 & \text{otherwise}\,, \end{cases} \tag{11}$$

where $|\mathcal{M}|$ is the determinant of $\mathcal{M}_{\mu\nu}$ and c_* is an *ad-hoc* cutoff[1] used to normalize the prior, namely $C = \frac{\sqrt{|\mathcal{M}|}}{(2\pi)^2}\,c_*^{-1}$.

Using this prior and taking the integration boundary to infinity, $c_* \to \infty$, we obtain the (marginal) signal likelihood, from Eq. (10), in the form

$$P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{S}\right) = P\left(\boldsymbol{x}|\mathcal{H}_\mathrm{G}\right)\,c_*^{-1}\,\mathrm{e}^{\mathcal{F}(\boldsymbol{x})}\,, \tag{12}$$

where we define the (coherent) multi-detector $\mathcal{F}$-statistic as

$$2\mathcal{F}(\boldsymbol{x}) \equiv x_\mu\,\mathcal{M}^{\mu\nu}\,x_\nu\,, \tag{13}$$

and $\mathcal{M}^{\mu\nu}$ denotes the inverse matrix to $\mathcal{M}_{\mu\nu}$, i.e., $\mathcal{M}_{\mu\alpha}\mathcal{M}^{\alpha\nu} = \delta_\mu^\nu$. We obtain the posterior probability for the signal hypothesis as

$$P\left(\mathcal{H}_\mathrm{S}|\boldsymbol{x}\right) = o_\mathrm{SG}\,c_*^{-1}\,P\left(\mathcal{H}_\mathrm{G}|\boldsymbol{x}\right)\,\mathrm{e}^{\mathcal{F}(\boldsymbol{x})}\,, \tag{14}$$

---

[1] This translates to the notation of Ref. [35] via $c_* = \frac{\hat{\rho}_{\max}^4}{70}$.

where $o_{\mathrm{SG}} \equiv P\left(\mathcal{H}_{\mathrm{S}}\right)/P\left(\mathcal{H}_{\mathrm{G}}\right)$ denotes the prior odds between the signal- and Gaussian-noise hypotheses.

The posterior odds between signal hypothesis $\mathcal{H}_{\mathrm{S}}$ and Gaussian-noise hypothesis $\mathcal{H}_{\mathrm{G}}$ are therefore equivalent to the standard multi-detector $\mathcal{F}$-statistic[2], as we see by writing

$$O_{\mathrm{SG}}(\boldsymbol{x}) \equiv \frac{P\left(\mathcal{H}_{\mathrm{S}}|\boldsymbol{x}\right)}{P\left(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x}\right)} = o_{\mathrm{SG}}\, c_*^{-1}\, \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}\,. \qquad (15)$$

Note that the corresponding (marginal) likelihood ratio

$$B_{\mathrm{SG}}(\boldsymbol{x}) \equiv \frac{P\left(\boldsymbol{x}|\mathcal{H}_{\mathrm{S}}\right)}{P\left(\boldsymbol{x}|\mathcal{H}_{\mathrm{G}}\right)} = c_*^{-1}\, \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}\,, \qquad (16)$$

is generally known as the *Bayes factor*, and is closely related to the odds via $O_{\mathrm{SG}}(\boldsymbol{x}) = o_{\mathrm{SG}}\, B_{\mathrm{SG}}(\boldsymbol{x})$.

While this statistic is close to optimal for detecting signals in pure Gaussian noise [6], it is vulnerable to various signal-like instrumental artifacts in the data. As discussed in Sec. I, we see from Eqs. (15) and (16) that detector artifacts can trigger $O_{\mathrm{SG}}(\boldsymbol{x})$ or $B_{\mathrm{SG}}(\boldsymbol{x})$, provided they resemble $\mathcal{H}_{\mathrm{S}}$ *more* than $\mathcal{H}_{\mathrm{G}}$ even if the agreement with $\mathcal{H}_{\mathrm{S}}$ is poor. In order to deal with this problem, we need to introduce an alternative hypothesis, which describes instrumental lines *better* than $\mathcal{H}_{\mathrm{S}}$.

C. Simple line hypothesis: a CW-like disturbance in a single detector

Here we introduce a simple line hypothesis designed to match one prominent feature of many instrumental lines, distinguishing them from CW signals: the fact that they appear only in one detector. Inspired by this, we re-use the signal hypothesis from Eq. (5) in order to define a line in detector X :

$$\begin{aligned} \mathcal{H}_{\mathrm{L}}^{X} : x^{X}(t) &= n^{X}(t) + h^{X}(t; \mathcal{A}^{X}) \\ &\text{with prior} \quad P\left(\mathcal{A}^{X}|\mathcal{H}_{\mathrm{L}}^{X}\right)\,. \end{aligned} \qquad (17)$$

We would expect lines to have a different amplitude distribution than real signals, but in the absence of any more detailed knowledge on this point, we choose to re-use the signal amplitude prior given by Eq. (11) for $P\left(\mathcal{A}^{X}|\mathcal{H}_{\mathrm{L}}^{X}\right)$. This choice simplifies the following calculations. In analogy to Eq. (14), we directly obtain the probability for $\mathcal{H}_{\mathrm{L}}^{X}$:

$$P\left(\mathcal{H}_{\mathrm{L}}^{X}|x^{X}\right) = c_*^{-1}\, P\left(\mathcal{H}_{\mathrm{G}}^{X}|x^{X}\right)\, o_{\mathrm{LG}}^{X}\, \mathrm{e}^{\mathcal{F}^{X}(x^{X})}\,. \qquad (18)$$

Here we define the per-detector prior line odds $o_{\mathrm{LG}}^{X} \equiv P\left(\mathcal{H}_{\mathrm{L}}^{X}\right)/P\left(\mathcal{H}_{\mathrm{G}}^{X}\right)$, which encode prior knowledge about how likely a line is, compared to pure Gaussian noise, in

a given template λ and detector X. The detector-specific $\mathcal{F}$-statistic $\mathcal{F}^{X}(x^{X})$ is simply given by Eq. (13) restricted to detector X.

For multiple detectors we can now formulate the simple line hypothesis $\mathcal{H}_{\mathrm{L}}$ as a CW-like disturbance $\mathcal{H}_{\mathrm{L}}^{X}$ in any one detector X and data consistent with Gaussian noise $\mathcal{H}_{\mathrm{G}}^{Y}$ in all other detectors $Y \neq X$:

$$\begin{aligned} \mathcal{H}_{\mathrm{L}} \equiv &\left(\mathcal{H}_{\mathrm{L}}^{1} \text{ and } \mathcal{H}_{\mathrm{G}}^{2} \text{ and } \mathcal{H}_{\mathrm{G}}^{3} \ldots\right) \text{ or} \\ &\left(\mathcal{H}_{\mathrm{G}}^{1} \text{ and } \mathcal{H}_{\mathrm{L}}^{2} \text{ and } \mathcal{H}_{\mathrm{G}}^{3} \ldots\right) \text{ or } \ldots\,. \end{aligned} \qquad (19)$$

Note that in this approach $\mathcal{H}_{\mathrm{L}}$ does not include lines that are coincident across different detectors, which is postponed to future work.

We assume the different detectors to be independent to the extent that knowing $\mathcal{H}_{\mathrm{G}}^{X}$ or $\mathcal{H}_{\mathrm{L}}^{X}$ for detector X does not inform us about $\mathcal{H}_{\mathrm{G}}^{Y}$ or $\mathcal{H}_{\mathrm{L}}^{Y}$ for other detectors $Y \neq X$. We also assume the different alternatives in Eq. (19) to be mutually exclusive. The laws of probability therefore yield

$$\begin{aligned} P\left(\mathcal{H}_{\mathrm{L}}|\boldsymbol{x}\right) &= P\left(\mathcal{H}_{\mathrm{L}}^{1}|x^{1}\right) P\left(\mathcal{H}_{\mathrm{G}}^{2}|x^{2}\right) P\left(\mathcal{H}_{\mathrm{G}}^{3}|x^{3}\right) \times \ldots \\ &+ P\left(\mathcal{H}_{\mathrm{G}}^{1}|x^{1}\right) P\left(\mathcal{H}_{\mathrm{L}}^{2}|x^{2}\right) P\left(\mathcal{H}_{\mathrm{G}}^{3}|x^{3}\right) \times \ldots \\ &+ \ldots \\ &= \sum_{X} P\left(\mathcal{H}_{\mathrm{L}}^{X}|x^{X}\right) \prod_{Y \neq X} P\left(\mathcal{H}_{\mathrm{G}}^{Y}|x^{Y}\right)\,. \end{aligned} \qquad (20)$$

By combining Eqs. (19), (18) and the (per-detector) Gaussian-noise probability from Eq. (4), we find the posterior probability for the line hypothesis $\mathcal{H}_{\mathrm{L}}$ as

$$P\left(\mathcal{H}_{\mathrm{L}}|\boldsymbol{x}\right) = c_*^{-1}\, P\left(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x}\right) \sum_{X} o_{\mathrm{LG}}^{X}\, \mathrm{e}^{\mathcal{F}^{X}(x^{X})}\,, \qquad (21)$$

where we used the fact that $\prod_{X} P\left(\mathcal{H}_{\mathrm{G}}^{X}|x^{X}\right) = P\left(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x}\right)$. Note that

$$\sum_{X} o_{\mathrm{LG}}^{X} = \frac{P\left(\mathcal{H}_{\mathrm{L}}\right)}{P\left(\mathcal{H}_{\mathrm{G}}\right)} \equiv o_{\mathrm{LG}}\,, \qquad (22)$$

where o_{LG} denotes the prior odds for a line versus Gaussian noise (in the present template λ) including all detectors.

It will be convenient to define relative detector weights r^{X} for the prior line odds, namely for N_{det} detectors:

$$r^{X} \equiv \frac{o_{\mathrm{LG}}^{X}}{o_{\mathrm{LG}}/N_{\mathrm{det}}}\,, \quad \text{such that} \quad \sum_{X} r^{X} = N_{\mathrm{det}}\,. \qquad (23)$$

If all detectors are equally likely to contain a line, then $r^{X} = 1$ for all X. We further denote the average of a quantity Q^{X} over detectors as

$$\left\langle Q^{X} \right\rangle \equiv \frac{1}{N_{\mathrm{det}}} \sum_{X} Q^{X}\,, \qquad (24)$$

and hence $\left\langle r^{X} \right\rangle = 1$. By using these definitions, we can write Eq. (21) as follows:

$$P\left(\mathcal{H}_{\mathrm{L}}|\boldsymbol{x}\right) = c_*^{-1}\, P\left(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x}\right)\, o_{\mathrm{LG}} \left\langle r^{X}\, \mathrm{e}^{\mathcal{F}^{X}(x^{X})} \right\rangle\,. \qquad (25)$$

---

[2] *Equivalence* in the Neyman-Pearson sense: the same false-dismissal as a function of false-alarm probability.

IV. COHERENT LINE-ROBUST STATISTICS

We use the posterior line probability of Eq. (25) to compute the odds for additional model comparisons, thereby extending the standard multi-detector $\mathcal{F}$-statistic given by Eq. (15). In particular, we consider two approaches:

- Define a "line-veto" statistic as the odds between the signal hypothesis $\mathcal{H}_\mathrm{S}$ and the line hypothesis $\mathcal{H}_\mathrm{L}$. This may be useful, for example, as a follow-up statistic for strong candidates from an initial $\mathcal{F}$-statistic search, which compared $\mathcal{H}_\mathrm{S}$ versus Gaussian noise $\mathcal{H}_\mathrm{G}$. In such a two-stage approach, one would test the signal hypothesis against the line-hypothesis if the Gaussian-noise hypothesis has been ruled out with sufficient confidence.

- *Extend* the standard signal-versus-Gaussian-noise odds $O_\mathrm{SG}(\boldsymbol{x})$ to a more line-robust statistic $O_\mathrm{SGL}(\boldsymbol{x})$ by allowing the noise hypothesis to include either pure Gaussian noise $\mathcal{H}_\mathrm{G}$ or a line $\mathcal{H}_\mathrm{L}$.

A. Line-veto statistic $O_\mathrm{SL}(\boldsymbol{x})$

Using the posterior probabilities given by Eqs. (14) and (25), we obtain the posterior signal-versus-line odds as

$$O_\mathrm{SL}(\boldsymbol{x}) \equiv \frac{P\left(\mathcal{H}_\mathrm{S}|\boldsymbol{x}\right)}{P\left(\mathcal{H}_\mathrm{L}|\boldsymbol{x}\right)} = o_\mathrm{SL} \; \frac{\mathrm{e}^{\mathcal{F}(\boldsymbol{x})}}{\left\langle r^X \mathrm{e}^{\mathcal{F}^X(x^X)}\right\rangle}, \qquad (26)$$

with the prior odds $o_\mathrm{SL} \equiv P\left(\mathcal{H}_\mathrm{S}\right)/P\left(\mathcal{H}_\mathrm{L}\right) = o_\mathrm{SG}/o_\mathrm{LG}$. Note that the amplitude-prior cutoff c_* has disappeared, as we have used the same amplitude prior on lines and signals.

In the following we will often neglect the dependency on $\boldsymbol{x}$ and x^X to simplify notation. It is instructive to consider the log-odds, which we can write as

$$\begin{aligned} \ln O_\mathrm{SL} &= \ln o_\mathrm{SL} + \mathcal{F} - \ln\left\langle r^X \mathrm{e}^{\mathcal{F}^X}\right\rangle \\ &= \ln o_\mathrm{SL} + \mathcal{F} - \mathcal{F}'_\mathrm{max} - \ln\left\langle r^X \mathrm{e}^{\left(\mathcal{F}^X - \mathcal{F}'_\mathrm{max}\right)}\right\rangle, \end{aligned} \tag{27}$$

where we define

$$\mathcal{F}'_\mathrm{max} \equiv \max_X \left(\mathcal{F}^X + \ln r^X\right). \tag{28}$$

The terms in the detector-average in Eq. (27) are bounded within $[0,1]$, with at least one term being equal to 1, and therefore the logarithmic average $\ln\langle\ldots\rangle$ is bounded within $[-\ln N_\mathrm{det}, 0] \sim \mathcal{O}(1)$. For strong $\mathcal{F}$-statistic candidates, i.e., $\mathcal{F} \gg 1$, the logarithmic correction is negligible, and therefore we can approximate

$$\ln O_\mathrm{SL}(\boldsymbol{x}) \approx \mathcal{F}(\boldsymbol{x}) - \mathcal{F}'_\mathrm{max}(\boldsymbol{x}) + \ln o_\mathrm{SL}. \tag{29}$$

Without prior knowledge about one detector being more affected by instrumental lines than others, we would have

$r^X = 1$ and therefore $\mathcal{F}'_\mathrm{max}(x) = \max_X \mathcal{F}^X(x^X) \equiv \mathcal{F}_\mathrm{max}(x)$. Considered as a detection statistic, $O_\mathrm{SL}(\boldsymbol{x})$ is therefore approximately equivalent to the difference between the multi-detector $\mathcal{F}$-statistic and the largest $\mathcal{F}$-statistic value from the individual detectors.

By choosing a special threshold of $O_\mathrm{SL}(\boldsymbol{x}) = o_\mathrm{SL}$ and assuming equal prior line probabilities for all detectors, we recover the well-known $\mathcal{F}$-*statistic consistency veto*, namely

$$\text{If} \quad \mathcal{F}(\boldsymbol{x}) < \max_X \{\mathcal{F}^X(x)\} \quad \Longrightarrow \quad \text{veto the candidate}, \tag{30}$$

which has been successfully used and tested in Refs. [11, 28, 32]. Combining this veto with $\mathcal{F}$-statistic ranking corresponds to defining a new statistic:

$$\mathcal{F}^{+\mathrm{veto}}(\boldsymbol{x}) \equiv \begin{cases} \mathcal{F}(\boldsymbol{x}) & \text{if } \mathcal{F}(\boldsymbol{x}) \geq \max_X \{\mathcal{F}^X(x)\}, \\ 0 & \text{otherwise}, \end{cases} \tag{31}$$

which we will refer to as the $\mathcal{F}^{+\mathrm{veto}}$-statistic.

B. Line-robust detection statistic $O_\mathrm{SGL}(\boldsymbol{x})$

From the standpoint of probability theory it is more natural to use the line hypothesis to extend what we mean by "noise", namely either pure Gaussian noise $\mathcal{H}_\mathrm{G}$ or a line $\mathcal{H}_\mathrm{L}$. Hence, we introduce an extended noise hypothesis as

$$\mathcal{H}_\mathrm{GL} : \mathcal{H}_\mathrm{G} \text{ or } \mathcal{H}_\mathrm{L}. \tag{32}$$

Since we take $\mathcal{H}_\mathrm{G}$ and $\mathcal{H}_\mathrm{L}$ to be mutually exclusive, the posterior probability for $\mathcal{H}_\mathrm{GL}$ is

$$\begin{aligned} P\left(\mathcal{H}_\mathrm{GL}|\boldsymbol{x}\right) &= P\left(\mathcal{H}_\mathrm{G}|\boldsymbol{x}\right) + P\left(\mathcal{H}_\mathrm{L}|\boldsymbol{x}\right) \\ &= P\left(\mathcal{H}_\mathrm{G}|\boldsymbol{x}\right)\left(1 + c_*^{-1} o_\mathrm{LG} \left\langle r^X \mathrm{e}^{\mathcal{F}^X(x^X)}\right\rangle\right), \end{aligned} \tag{33}$$

where we have used Eq. (25) for the explicit line posterior.

Interestingly, we can express the odds $O_\mathrm{SGL}(\boldsymbol{x})$ of the signal versus extended noise hypotheses as

$$O_\mathrm{SGL}(\boldsymbol{x}) \equiv \frac{P\left(\mathcal{H}_\mathrm{S}|\boldsymbol{x}\right)}{P\left(\mathcal{H}_\mathrm{GL}|\boldsymbol{x}\right)} = \left[O_\mathrm{SG}^{-1}(\boldsymbol{x}) + O_\mathrm{SL}^{-1}(\boldsymbol{x})\right]^{-1}. \tag{34}$$

We can compare this result with the *ad-hoc* two-stage approach discussed previously, where one would set two independent thresholds on O_SG and on O_SL. As we see from Eq. (34), instead the laws of probability tell us to compute the harmonic sum of O_SG and O_SL and to set a single threshold on the resulting statistic.

Inserting the explicit expressions provided by Eqs. (33) and (14), we obtain

$$O_\mathrm{SGL}(\boldsymbol{x}) = \frac{o_\mathrm{SG} \, \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}}{c_* + o_\mathrm{LG} \left\langle r^X \mathrm{e}^{\mathcal{F}^X}\right\rangle}. \tag{35}$$

Note that contrary to the previous odds O_{SG} and O_{SL}, now the choice of amplitude-prior cutoff parameter c_* from Eq. (11) will affect the properties of the resulting statistic.

We can rewrite Eq. (35) by introducing prior odds $o_{\mathrm{SGL}} \equiv P(\mathcal{H}_{\mathrm{S}})/P(\mathcal{H}_{\mathrm{GL}})$ and (noting that $o_{\mathrm{SG}} = o_{\mathrm{SGL}}(1 + o_{\mathrm{LG}})$) we obtain

$$O_{\mathrm{SGL}}(\boldsymbol{x}) = o_{\mathrm{SGL}} \frac{\mathrm{e}^{\mathcal{F}(\boldsymbol{x})}}{(1 - p_{\mathrm{L}})\,\mathrm{e}^{\mathcal{F}_*^{(0)}} + p_{\mathrm{L}}\,\left\langle r^X \mathrm{e}^{\mathcal{F}^X(x^X)}\right\rangle}, \tag{36}$$

where we define the prior line probability p_{L} as

$$p_{\mathrm{L}} \equiv \frac{o_{\mathrm{LG}}}{1 + o_{\mathrm{LG}}} = \frac{P(\mathcal{H}_{\mathrm{L}})}{P(\mathcal{H}_{\mathrm{GL}})} = P(\mathcal{H}_{\mathrm{L}}|\mathcal{H}_{\mathrm{GL}}) \in [0, 1]\,, \tag{37}$$

and we used a more natural re-parametrization of c_* by defining

$$\mathcal{F}_*^{(0)} \equiv \ln c_*\,. \tag{38}$$

Limiting cases of $O_{\mathrm{SGL}}(\boldsymbol{x})$

We now consider the limiting behavior of O_{SGL} as a function of the line prior p_{L} and of the single-detector $\mathcal{F}^X(x)$ values. We see from Eq. (36) that $O_{\mathrm{SGL}}(\boldsymbol{x})$ reduces to the $\mathcal{F}$-statistic if we are certain that there are no lines, i.e., $O_{\mathrm{SGL}}(\boldsymbol{x}) \to O_{\mathrm{SG}}(\boldsymbol{x}) \propto \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}$ for $p_{\mathrm{L}} \to 0$. On the other hand, it reduces to the pure line-veto statistic of Eq. 26 when we believe the noise to be completely dominated by lines, i.e., $O_{\mathrm{SGL}}(\boldsymbol{x}) \to O_{\mathrm{SL}}(\boldsymbol{x})$ for $p_{\mathrm{L}} \to 1$.

For fixed p_{L} we see that the transition between these two extremes depends on the $\mathcal{F}^X(x)$ values compared to the prior scale $\mathcal{F}_*^{(0)}$. To illustrate this more clearly, we first rewrite Eq. (36) using the relations $o_{\mathrm{SGL}} = p_{\mathrm{L}}\,o_{\mathrm{SL}}$ and $(1 - p_{\mathrm{L}})/p_{\mathrm{L}} = o_{\mathrm{LG}}^{-1}$. Introducing the "transition scale" $\mathcal{F}_*$ as

$$\mathcal{F}_* \equiv \mathcal{F}_*^{(0)} - \ln o_{\mathrm{LG}}\,, \tag{39}$$

we obtain

$$O_{\mathrm{SGL}}(\boldsymbol{x}) = o_{\mathrm{SL}} \frac{\mathrm{e}^{\mathcal{F}(\boldsymbol{x})}}{\mathrm{e}^{\mathcal{F}_*} + \left\langle r^X \mathrm{e}^{\mathcal{F}^X(x^X)}\right\rangle}\,. \tag{40}$$

From this reparametrization, we see that $\mathcal{F}_*$ defines the scale of a smooth transition of $O_{\mathrm{SGL}}(\boldsymbol{x})$ between $O_{\mathrm{SG}}(\boldsymbol{x}) \propto \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}$ and $O_{\mathrm{SL}}(\boldsymbol{x})$ depending on the values of $\mathcal{F}^X$: namely the "line-veto term" $\left\langle r^X \mathrm{e}^{\mathcal{F}^X}\right\rangle$ in Eq. (40) only starts to play a role when it is comparable to $\mathrm{e}^{\mathcal{F}_*}$.

To see this more explicitly, we write the log-odds as

$$\begin{aligned} \ln O_{\mathrm{SGL}}(\boldsymbol{x}) = {}& \ln o_{\mathrm{SL}} + \mathcal{F}(\boldsymbol{x}) - \mathcal{F}_{\mathrm{max}}''(x) \\ & - \ln\left(\mathrm{e}^{\mathcal{F}_* - \mathcal{F}_{\mathrm{max}}''} + \left\langle r^X \mathrm{e}^{\mathcal{F}^X(x^X) - \mathcal{F}_{\mathrm{max}}''(\boldsymbol{x})}\right\rangle\right), \end{aligned} \tag{41}$$

where we define $\mathcal{F}_{\mathrm{max}}''(\boldsymbol{x}) \equiv \max\left(\mathcal{F}_*,\,\mathcal{F}^X(x^X) + \ln r^X\right)$. The logarithmic correction is of order unity, therefore this

effectively corresponds to $\ln O_{\mathrm{SL}}(\boldsymbol{x})$ when $\max(\mathcal{F}^X(x) + \ln r^X) > \mathcal{F}_*$, and to $\ln O_{\mathrm{SG}}(\boldsymbol{x})$ otherwise.

In practice it can be difficult to determine good prior values for $\mathcal{F}_*^{(0)}$, due to the unphysical choice of amplitude priors in Eq. (11). We will discuss this issue in more detail in Sec. VI B.

This transitioning behavior is reminiscent of the two-stage line-veto approach discussed in Sec. IV. There one applies a line veto only to candidates that are "strong" in terms of $O_{\mathrm{SG}}(\boldsymbol{x}) \propto \mathrm{e}^{\mathcal{F}(\boldsymbol{x})}$, which means that the Gaussian-noise hypothesis is already considered sufficiently unlikely. Note, however, that for $O_{\mathrm{SGL}}(\boldsymbol{x})$ the transition from $O_{\mathrm{SG}}(\boldsymbol{x})$ to $O_{\mathrm{SL}}(\boldsymbol{x})$ is smooth and depends on the strength of the single-detector statistics $\mathcal{F}^X(x^X)$ rather than the multi-detector statistic $\mathcal{F}(\boldsymbol{x})$.

V. SEMI-COHERENT LINE-ROBUST STATISTICS

For unknown signal parameters λ, the use of the fully coherent (in time) $\mathcal{F}$-statistic is usually prohibitive in terms of computing cost. Thus, *semi-coherent* methods are typically used, being more sensitive at fixed computing cost [36, 37]. In this approach the data $\boldsymbol{x}$ is divided into N_{seg} segments of shorter duration, denoted as $\{\boldsymbol{x}_k\}_{k=1}^{N_{\mathrm{seg}}}$. The coherent statistic $\mathcal{F}_k(\boldsymbol{x}_k; \lambda)$ in a template λ is computed for each segment k separately and then combined *incoherently*, typically by summing over all data segments. This is often referred to as the "Stack-Slide" method. Other incoherent combinations such as the "Hough transform" method [38] will not be discussed here. The following discussion refers to the statistic in a single template λ, and we will therefore simplify the notation again by dropping λ.

As shown in Ref. [35], the semi-coherent StackSlide $\mathcal{F}$-statistic can be derived by relaxing the requirement of consistent signal amplitudes $\mathcal{A}$ across different segments, i.e., allowing for a set of N_{seg} independent amplitude parameters $\mathcal{A}_k$ in Eq. (5). This defines the semi-coherent signal hypothesis $\widehat{\mathcal{H}}_{\mathrm{S}}$ as

$$\widehat{\mathcal{H}}_{\mathrm{S}} : \boldsymbol{x}_k = \boldsymbol{n}_k + \mathcal{A}_k^\mu\,\boldsymbol{h}_\mu\,, \quad \text{for } k = 1, \ldots N_{\mathrm{seg}}\,, \tag{42}$$

where here and in the following the hat $\widehat{}$ notation refers to semi-coherent quantities.

For the per-segment amplitude priors $P\left(\mathcal{A}_k|\widehat{\mathcal{H}}_{\mathrm{S}}\right)$, we reuse the amplitude prior given by Eq. (11). Hence, by marginalization as in Eq. (10), we obtain the posterior

$$P\left(\widehat{\mathcal{H}}_{\mathrm{S}}|\boldsymbol{x}\right) = \widehat{o}_{\mathrm{SG}}\,P(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x})\,c_*^{-N_{\mathrm{seg}}}\,\mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}\,, \tag{43}$$

where we define the StackSlide $\mathcal{F}$-statistic $\widehat{\mathcal{F}}$ in parameter-space point λ as

$$\widehat{\mathcal{F}}(\boldsymbol{x}; \lambda) \equiv \sum_{k=1}^{N_{\mathrm{seg}}} \mathcal{F}_k(\boldsymbol{x}_k; \lambda)\,. \tag{44}$$

For Gaussian noise we have $\widehat{\mathcal{H}}_{\mathrm{G}} = \mathcal{H}_{\mathrm{G}}$, but for consistency of notation we still write $\widehat{\mathcal{H}}_{\mathrm{G}}$ throughout this section. The posterior odds between the signal and Gaussian-noise hypotheses across the N_{seg} segments is

$$\widehat{O}_{\mathrm{SG}}(\boldsymbol{x}) \equiv \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}}|\boldsymbol{x}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{G}}|\boldsymbol{x}\right)} = \widehat{o}_{\mathrm{SG}}\, c_*^{-N_{\mathrm{seg}}}\, \mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}. \tag{45}$$

We can now generalize the single-detector line hypothesis of Eq. (19) to the semi-coherent case as done for the signal hypothesis in Eq. (42), namely:

$$\widehat{\mathcal{H}}_{\mathrm{L}} = \left(\widehat{\mathcal{H}}_{\mathrm{L}}^1 \text{ and } \widehat{\mathcal{H}}_{\mathrm{G}}^2 \text{ and } \widehat{\mathcal{H}}_{\mathrm{G}}^3 \dots\right) \text{ or} \\ \left(\widehat{\mathcal{H}}_{\mathrm{G}}^1 \text{ and } \widehat{\mathcal{H}}_{\mathrm{L}}^2 \text{ and } \widehat{\mathcal{H}}_{\mathrm{G}}^3 \dots\right) \text{ or } \dots \tag{46}$$

The probability of the line hypothesis in detector X across all segments is:

$$P\left(\widehat{\mathcal{H}}_{\mathrm{L}}^X|x^X\right) = P\left(\widehat{\mathcal{H}}_{\mathrm{G}}^X|x^X\right) c_*^{-N_{\mathrm{seg}}}\, \widehat{o}_{\mathrm{LG}}^X\, \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)}, \tag{47}$$

where the semi-coherent line-odds in detector X is $\widehat{o}_{\mathrm{LG}}^X \equiv P\left(\widehat{\mathcal{H}}_{\mathrm{L}}^X\right)/P\left(\widehat{\mathcal{H}}_{\mathrm{G}}^X\right)$. Similarly to Eq. (25), the posterior probability for the semi-coherent line-hypothesis $\widehat{\mathcal{H}}_{\mathrm{L}}$ is obtained as

$$P\left(\widehat{\mathcal{H}}_{\mathrm{L}}|\boldsymbol{x}\right) = P\left(\widehat{\mathcal{H}}_{\mathrm{G}}|\boldsymbol{x}\right) c_*^{-N_{\mathrm{seg}}}\, \widehat{o}_{\mathrm{LG}} \left\langle \widehat{r}^X\, \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)} \right\rangle, \tag{48}$$

where in analogy to Eqs. (22) and (23) we define

$$\widehat{o}_{\mathrm{LG}} \equiv \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{L}}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{G}}\right)} = \sum_X \widehat{o}_{\mathrm{LG}}^X, \tag{49}$$

$$\widehat{r}^X \equiv \frac{\widehat{o}_{\mathrm{LG}}^X}{\widehat{o}_{\mathrm{LG}}/N_{\mathrm{det}}}. \tag{50}$$

The posterior probability for the extended noise hypothesis,

$$\widehat{\mathcal{H}}_{\mathrm{GL}} \equiv \widehat{\mathcal{H}}_{\mathrm{G}} \text{ or } \widehat{\mathcal{H}}_{\mathrm{L}}, \tag{51}$$

is therefore given by

$$P\left(\widehat{\mathcal{H}}_{\mathrm{GL}}|\boldsymbol{x}\right) = P\left(\mathcal{H}_{\mathrm{G}}|\boldsymbol{x}\right) \left(1 + c_*^{-N_{\mathrm{seg}}}\, \widehat{o}_{\mathrm{LG}} \left\langle \widehat{r}^X\, \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)} \right\rangle\right). \tag{52}$$

We can now define a semi-coherent line-veto statistic, namely

$$\widehat{O}_{\mathrm{SL}}(\boldsymbol{x}) \equiv \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}}|\boldsymbol{x}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{L}}|\boldsymbol{x}\right)} = \widehat{o}_{\mathrm{SL}} \frac{\mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}}{\left\langle \widehat{r}^X\, \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)} \right\rangle}, \tag{53}$$

and a semi-coherent line-robust detection statistic as

$$\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x}) \equiv \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}}|\boldsymbol{x}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{GL}}|\boldsymbol{x}\right)} = \left[\widehat{O}_{\mathrm{SG}}^{-1}(\boldsymbol{x}) + \widehat{O}_{\mathrm{SL}}^{-1}(\boldsymbol{x})\right]^{-1}. \tag{54}$$

The latter can be written explicitly as

$$\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x}) = \widehat{o}_{\mathrm{SGL}} \frac{\mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}}{\left(1 - \widehat{p}_{\mathrm{L}}\right) \mathrm{e}^{\widehat{\mathcal{F}}_*^{(0)}} + \widehat{p}_{\mathrm{L}} \left\langle \widehat{r}^X \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)} \right\rangle}, \tag{55}$$

with (semi-coherent) line probability

$$\widehat{p}_{\mathrm{L}} \equiv \frac{\widehat{o}_{\mathrm{LG}}}{1 + \widehat{o}_{\mathrm{LG}}} = P\left(\widehat{\mathcal{H}}_{\mathrm{L}}|\widehat{\mathcal{H}}_{\mathrm{GL}}\right) \in [0, 1] \tag{56}$$

and, in analogy to Eq. (38), a prior cutoff parametrization of

$$\widehat{\mathcal{F}}_*^{(0)} \equiv \ln c_*^{N_{\mathrm{seg}}}. \tag{57}$$

Similarly to Eq. (40), we can therefore write this equivalently as

$$\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x}) = \widehat{o}_{\mathrm{SL}} \frac{\mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}}{\mathrm{e}^{\widehat{\mathcal{F}}_*} + \left\langle \widehat{r}^X \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X)} \right\rangle}, \tag{58}$$

where the semi-coherent transition scale $\widehat{\mathcal{F}}_*$ is defined as

$$\widehat{\mathcal{F}}_* \equiv \widehat{\mathcal{F}}_*^{(0)} - \ln \widehat{o}_{\mathrm{LG}}, \tag{59}$$

by generalizing Eq. (39). Hence, we find that $\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x})$ transitions from the standard semi-coherent statistic $\widehat{O}_{\mathrm{SG}}(\boldsymbol{x}) \propto \mathrm{e}^{\widehat{\mathcal{F}}}$ to the line-veto statistic $\widehat{O}_{\mathrm{SL}}(\boldsymbol{x})$ when

$$\left\langle \widehat{r}^X \mathrm{e}^{\widehat{\mathcal{F}}^X} \right\rangle \sim \mathrm{e}^{\widehat{\mathcal{F}}_*}. \tag{60}$$

We can rewrite the log-odds as

$$\ln \widehat{O}_{\mathrm{SGL}}(\boldsymbol{x}) = \ln \widehat{o}_{\mathrm{SL}} + \widehat{\mathcal{F}}(\boldsymbol{x}) - \widehat{\mathcal{F}}_{\mathrm{max}}''(\boldsymbol{x}) \\ - \ln \left(\mathrm{e}^{\widehat{\mathcal{F}}_* - \widehat{\mathcal{F}}_{\mathrm{max}}''(\boldsymbol{x})} + \left\langle \widehat{r}^X \mathrm{e}^{\widehat{\mathcal{F}}^X(x^X) - \widehat{\mathcal{F}}_{\mathrm{max}}''(\boldsymbol{x})} \right\rangle\right), \tag{61}$$

with $\widehat{\mathcal{F}}_{\mathrm{max}}''(\boldsymbol{x}) \equiv \max\left(\widehat{\mathcal{F}}_*, \widehat{\mathcal{F}}^X(x^X) + \ln \widehat{r}^X\right)$. Note that in the semi-coherent case we typically deal with much larger numerical values of $\widehat{\mathcal{F}}$ (due to its definition as a sum over segments in Eq. (44)). However, the logarithmic correction term is still of order unity. This implies that the transition from $\widehat{O}_{\mathrm{SG}}(\boldsymbol{x})$ to the line-veto odds $\widehat{O}_{\mathrm{SL}}(\boldsymbol{x})$ is expected to be sharper than in the coherent case of Eq. (41).

Incorporating the *ad-hoc* $\mathcal{F}$-statistic consistency veto discussed in Sec. IV A, we can define a semi-coherent $\widehat{\mathcal{F}}^{+\mathrm{veto}}$-statistic as

$$\widehat{\mathcal{F}}^{+\mathrm{veto}}(\boldsymbol{x}) \equiv \begin{cases} \widehat{\mathcal{F}}(\boldsymbol{x}) & \text{if } \widehat{\mathcal{F}}(\boldsymbol{x}) \geq \max_X\{\widehat{\mathcal{F}}^X(x)\}, \\ 0 & \text{otherwise}. \end{cases} \tag{62}$$

VI. CHOICE OF PRIORS

The new line-veto and line-robust statistics derived in this paper depend on some prior parameters which need

to be specified. We will now discuss a way to set their values.

The coherent statistics described in Sec. IV are simply special cases of the semi-coherent expressions given in Sec. V for $N_{\mathrm{seg}} = 1$. Hence, in the following we can use the semi-coherent notation without loss of generality.

In the pure line-veto statistic $\widehat{O}_{\mathrm{SL}}(\boldsymbol{x})$ of Eq. (53), the only relevant free parameters are the N_{det} line-probability weights $\widehat{r}^X$. This is because all monotonic functions of a test statistic are equivalent in the Neyman-Pearson sense, and the prior odds $\widehat{o}_{\mathrm{SL}}$ represent a proportionality factor only.

The line-robust statistic $\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x})$ depends on the prior odds $\widehat{o}_{\mathrm{LG}}$ and on the amplitude-prior cutoff parameter c_*. However, these two prior parameters only appear in $\widehat{O}_{\mathrm{SGL}}$ through the combination $\widehat{\mathcal{F}}_* \equiv \widehat{\mathcal{F}}_*^{(0)} - \ln \widehat{o}_{\mathrm{LG}}$ as defined in Eq. (59). Therefore $\widehat{O}_{\mathrm{SGL}}$ has $N_{\mathrm{det}} + 1$ free parameters.

While the prior odds $\widehat{o}_{\mathrm{LG}}$ have a clear intuitive interpretation, this is not the case for the prior amplitude cutoff parameter c_* and thus for $\widehat{\mathcal{F}}_*^{(0)}$, as defined in Eq. (57). This parameter results from the rather unphysical choice of the amplitude prior in Eq. (11), as discussed in more detail in Refs. [6, 35]. Hence, a certain amount of empirical "tuning" will be required to determine a reasonable value for $\widehat{\mathcal{F}}_*^{(0)}$, which we will discuss in Sec. VI B.

A. Proxy estimate of prior line probabilities from the data

A maximally uninformative choice for the line-priors would be $\widehat{r}^X = 1$ and $\widehat{o}_{\mathrm{LG}} = 1$, where the presence of lines is considered just as likely as pure Gaussian noise and all detectors are equally likely to be affected by lines. A more informed choice should be based on prior characterization of the detectors.

A practical way to achieve this is to judiciously use the observed data $\boldsymbol{x}$ for a simple "proxy" estimate of $\widehat{o}_{\mathrm{LG}}^X$. Empirically we find promising results when adopting the line-flagging method of Ref. [39]. We use data from all frequency bins potentially contributing to the detection statistics in a given search band. We compute the time-averaged normalized power over these bins and count how many exceed a predetermined threshold. The measured fraction of such outliers is used as a proxy estimate for the prior line probability.

More specifically, the data for $\mathcal{F}$-statistic searches is usually prepared in the form of *Short Fourier Transforms* (SFTs) of the original time-domain data, conventionally spanning stretches of duration $T_{\mathrm{SFT}} = 1800\,\mathrm{s}$ (e.g., see Ref. [38]). We compute the normalized average SFT power $\mathcal{P}^X(f)$ for each detector X as (e.g., see Ref. [40])

$$\mathcal{P}^X(f) \equiv \frac{2}{N_{\mathrm{SFT}} \, T_{\mathrm{SFT}}} \sum_{\alpha=1}^{N_{\mathrm{SFT}}} \frac{\left|\widetilde{x}_\alpha^X(f)\right|^2}{S_{\mathrm{n}}^{X\alpha}(f)} \,, \tag{63}$$

where the sum is over all N_{SFT} SFTs, $\widetilde{x}_\alpha^X(f)$ and $S_{\mathrm{n}}^{X\alpha}(f)$

denote the Fourier-transformed data and the noise PSD in the α-th SFT, respectively.

We estimate the prior line probability $\widehat{p}_{\mathrm{L}}^X$ for that frequency band as

$$\widehat{p}_{\mathrm{L}}^X = \frac{N_{\mathcal{P}>\mathcal{P}_{\mathrm{thr}}}^X}{N_{\mathrm{bins}}} \,, \tag{64}$$

where $N_{\mathcal{P}>\mathcal{P}_{\mathrm{thr}}}^X$ is the number of bins $\in [0, N_{\mathrm{bins}}]$ for which $\mathcal{P}^X(f)$ crossed the threshold $\mathcal{P}_{\mathrm{thr}}^X$. A typical band is of the order of $100\,\mathrm{mHz}$ wide, corresponding to a few hundred bins. The threshold $\mathcal{P}_{\mathrm{thr}}^X$ is chosen empirically to be safely above the typical noise fluctuations in the data.

From $\widehat{p}_{\mathrm{L}}^X$ the prior line odds may be computed as

$$\widehat{o}_{\mathrm{LG}}^X = \frac{\widehat{p}_{\mathrm{L}}^X}{1 - \widehat{p}_{\mathrm{L}}^X} \,, \tag{65}$$

which also fully specifies $\widehat{r}^X$ and $\widehat{o}_{\mathrm{LG}}$ via Eqs. (49) and (50).

We determine the threshold $\mathcal{P}_{\mathrm{thr}}^X$ by fixing a certain false-alarm probability $p_{\mathrm{FA}, \mathcal{P}}$. For large N_{SFT} this can be computed approximately from a Gaussian distribution with unit mean and standard deviation $\sigma = 1/\sqrt{N_{\mathrm{SFT}}}$.

As an illustrative example, Fig. 1 shows $\mathcal{P}^X(f)$ for a $\sim 60\,\mathrm{mHz}$ wide band of simulated Gaussian data consisting of 50 SFTs. The data is generated with a noise PSD of $S_{\mathrm{n}}^X = 3 \times 10^{-22}\,\mathrm{Hz}^{-1/2}$ in two detectors $X \in \{\mathrm{H1}, \mathrm{L1}\}$, where H1 and L1 stand for the LIGO detectors at Hanford and Livingston, respectively. A monochromatic stationary line of amplitude $h_0 = 2 \times 10^{-23}\,\mathrm{Hz}^{-1/2}$ at $50\,\mathrm{Hz}$ is injected in H1 only. More examples from real data are presented in Sec. VII B.
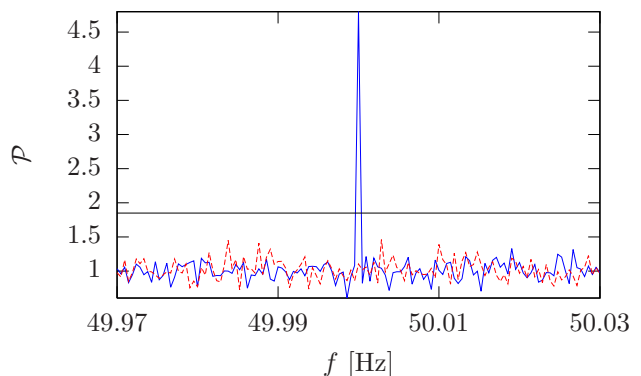

FIG. 1. Example of the normalized SFT power $\mathcal{P}^X(f)$ as a function of frequency f for LIGO H1 (*solid*) and L1 (*dashed*) for simulated Gaussian data containing a line in H1. The horizontal line shows the threshold $\mathcal{P}_{\mathrm{thr}}$ at a false-alarm level of $p_{\mathrm{FA}, \mathcal{P}} = 10^{-9}$.

We stress that the line-flagging procedure proposed here is not meant to yield a direct estimator of $\widehat{p}_{\mathrm{L}}^X$ but rather to provide an indication for the presence of lines based on spectral features that can be robustly identified.

For instance, observing no threshold crossings in the average SFT power $\mathcal{P}$ does not necessarily imply that the $\mathcal{F}$-statistic could not be affected by instrumental artifacts, while seeing many outliers in $\mathcal{P}$ does not always yield high values of $\mathcal{F}$. Hence we will not consider values of $\widehat{o}_{\text{LG}}^X$ that suggest more confidence than seems justifiable, and truncate its range to

$$\widehat{o}_{\text{LG}}^X \in [0.001, \ 1000] \,. \tag{66}$$

For the example simulated data set used in Fig. 1 we can detail the method as follows: there are 127 frequency bins in the band considered, and for the threshold $\mathcal{P}_{\text{thr}}^{\text{H1}} = \mathcal{P}_{\text{thr}}^{\text{L1}} = \mathcal{P}_{\text{thr}}(p_{\text{FA},\mathcal{P}} = 10^{-9}, N_{\text{SFT}} = 50) \approx 1.84$ there is a single crossing in H1 and none in L1. Hence, we estimate the line priors as $\widehat{o}_{\text{LG}}^{\text{H1}} = \max\left(0.001, \frac{1/127}{1-1/127}\right) \approx 0.008$ and $\widehat{o}_{\text{LG}}^{\text{L1}} = \max\left(0.001, \frac{0/127}{1-0/127}\right) = 0.001$.

We believe that this data-dependent prior estimation is not prone to the "sample re-use fallacy" [41]. The reason is that the proxy estimate for $\widehat{o}_{\text{LG}}^X$ is sufficiently *independent* from the posterior for the line hypothesis $\mathcal{H}_{\text{L}}$, as they are derived from data sets with effectively very little data in common. The line hypothesis $\mathcal{H}_{\text{L}}$ (being based on the signal hypothesis $\mathcal{H}_{\text{S}}$) describes a narrow-band signal, which in each half-hour SFT is confined to a few bins. In fact the current $\mathcal{F}$-statistic implementation [42] uses only 16 frequency bins per SFT to construct the detection statistic, and they are very heavily weighted towards a few central ones. On the other hand, the line-flagging prior estimate uses $\sim \mathcal{O}\,(100-200)$ frequency bins and each counts equally in the estimate. Furthermore, the results in Sec. VII B show that this procedure appears to be "safe" also in the presence of (injected) signals.

B. Empirical choice of transition scale $\widehat{\mathcal{F}}_*$

An additional free parameter in the line-robust statistic $\widehat{O}_{\text{SGL}}(\boldsymbol{x})$, as expressed in Eq. (58), is the transition scale $\widehat{\mathcal{F}}_* = \widehat{\mathcal{F}}_*^{(0)} - \ln \widehat{o}_{\text{LG}}$ of Eq. (59).

As discussed in Sec. IV B, $\widehat{\mathcal{F}}_*$ sets the scale (in terms of $\widehat{\mathcal{F}}^X$) for the transition of $\widehat{O}_{\text{SGL}}$ from the signal-versus-Gaussian-noise odds $\widehat{O}_{\text{SG}} \propto \mathrm{e}^{\widehat{\mathcal{F}}}$ (for $\widehat{\mathcal{F}}^X \ll \widehat{\mathcal{F}}_*$) to the signal-versus-line odds $\widehat{O}_{\text{SL}}$ (for $\widehat{\mathcal{F}}^X \gg \widehat{\mathcal{F}}_*$). Thus we can interpret $\widehat{\mathcal{F}}_*^{(0)}$ as the transition scale in the case of even prior odds, i.e., $\widehat{o}_{\text{LG}} = 1$, between the line and Gaussian-noise hypotheses. The effect of $\widehat{o}_{\text{LG}}$, which we estimate with the method described in the previous section, is to shift the transition scale up or down from this baseline, depending on whether prior knowledge gives lines lower or higher odds, respectively.

We can therefore interpret $\widehat{\mathcal{F}}_*^{(0)}$ as the level of $\widehat{\mathcal{F}}^X$ at which the associated false-alarm probability is low enough that we would start to doubt the Gaussian-noise hypothesis and would have to consider alternatives. Hence, we can express $\widehat{\mathcal{F}}_*^{(0)}$ in terms of this false-alarm probability, denoted as $p_{\text{FA}*}^{(0)}$:

$$\widehat{\mathcal{F}}_*^{(0)} = \widehat{\mathcal{F}}_*^{(0)}\left(p_{\text{FA}*}^{(0)}, \ N_{\text{seg}}\right). \tag{67}$$

In Gaussian noise, $\widehat{\mathcal{F}}^X$ follows a central χ^2-distribution with $4N_{\text{seg}}$ degrees of freedom.

On the one hand, we want $\widehat{\mathcal{F}}_*^{(0)}$ to be low enough ($p_{\text{FA}*}^{(0)}$ high enough) to suppress even weak lines, but not so low as to compromise the performance in Gaussian noise. When most of the data is approximately Gaussian (as is typically the case for CW searches), a reasonable choice is to use the lowest $\widehat{\mathcal{F}}_*^{(0)}$ (highest $p_{\text{FA}*}^{(0)}$) that does not yet adversely affect the detection power in Gaussian noise. In practice, we resort to an empirical choice of $p_{\text{FA}*}^{(0)}$ based on Monte-Carlo simulations on a small subset of the data.

VII. PERFORMANCE TESTS

Here we will discuss the detection efficiency of the statistics introduced in the previous sections for a population of signals embedded in different types of noise. In order to do this we use two different and somewhat complementary approaches: *(i)* fully "synthetic" simulations, which allow for efficient large-scale explorations under idealized conditions, and *(ii)* injections of simulated signals into LIGO S5 data containing instrumental artifacts.

We compare the performance of the following statistics (the second equation always refers to the corresponding semi-coherent version):

1. Standard multi-detector $\mathcal{F}$-statistic, Eqs. (13) and (44)

2. $\mathcal{F}^{+\text{veto}}$-statistic, Eqs. (31) and (62)

3. Line-veto statistic O_{SL}, Eqs. (26) and (53)

4. Line-robust statistic O_{SGL}, Eqs. (40) and (58)

In the case of the line-robust statistic O_{SGL} we use different transition scales $\mathcal{F}_*^{(0)}$ corresponding to false-alarm levels $p_{\text{FA}*}^{(0)}$, which we denote as

$$^{(-n)}O_{\text{SGL}}(\boldsymbol{x}) \equiv O_{\text{SGL}}(\boldsymbol{x}; \ p_{\text{FA}*}^{(0)} = 10^{-n}). \tag{68}$$

In the following tests we use $^{(-1)}O_{\text{SGL}}$, $^{(-3)}O_{\text{SGL}}$, and $^{(-6)}O_{\text{SGL}}$, corresponding to transition-scale false-alarm levels of $p_{\text{FA}*}^{(0)} = 10^{-1}, 10^{-3}, 10^{-6}$, respectively.

In order to assess the importance of the choice of prior line odds o_{LG}^X, we consider two cases:

- Uninformative priors, i.e., $o_{\text{LG}}^X = 1$ for all X: the corresponding "untuned" statistics are denoted as $O_{\text{SL}}^{(0)}$ and $^{(-n)}O_{\text{SGL}}^{(0)}$.

- Line priors o_{LG}^X using prior information on the line population: the corresponding "tuned" statistics are denoted as O_{SL} and $^{(-n)}O_{\text{SGL}}$, respectively.

A. Tests using synthetic draws

In this section, for simplicity, we consider only the co-herent case (cf. Sec. IV). Using the synthesizing approach described in Refs. [6, 35], one can directly generate random draws of the various statistics of interest for pure noise and for noise containing a signal.

The synthesizing method consists in generating random draws of the $\{x_\mu^X\}$ of Eq. (9) using their known (multi-variate) Gaussian distribution. From these we compute the $\mathcal{F}$- and $\mathcal{F}^X$-statistics from Eq. (13), O_{SL} from Eq. (26) and O_{SGL} from Eq. (40). In the following we refer to each draw of $\{x_\mu^X\}$ together with the resulting statistics as a *candidate*.

We generate the noise draws in such a way that a fraction f_{L} contains a line according to $\mathcal{H}_{\mathrm{L}}$ of Eq. (19), namely a CW signal in a single detector. The remaining fraction $1 - f_{\mathrm{L}}$ of noise draws follows the Gaussian-noise hypothesis $\mathcal{H}_{\mathrm{G}}$ of Eq. (2). In the following we refer to f_{L} as the *line contamination*.

From the noise draws we estimate for each statistic a threshold corresponding to a particular false-alarm probability p_{FA}. Applying this threshold to the signal candidates yields the detection probability $p_{\mathrm{det}}(p_{\mathrm{FA}})$ for each statistic at the false-alarm level p_{FA}. This is known as the *receiver operator characteristic* (ROC).

The strength of the injected signals is characterized by the (multi-detector) *signal-to-noise ratio* ρ_{S}, defined in the usual way [4] as

$$\rho_{\mathrm{S}}^2 \equiv (h|h) = \mathcal{A}^\mu \mathcal{M}_{\mu\nu} \mathcal{A}^\nu . \qquad (69)$$

This is related to the expectation value of the $\mathcal{F}$-statistic as $E[2\mathcal{F}]_{\mathcal{H}_{\mathrm{S}}} = 4 + \rho_{\mathrm{S}}^2$. As shown in Appendix A, for a line according to $\mathcal{H}_{\mathrm{L}}$ in detector Y, the expectation value of the multi-detector $\mathcal{F}$-statistic is approximately

$$E[2\mathcal{F}]_{\mathcal{H}_{\mathrm{L}}} \approx 4 + \frac{1}{N_{\mathrm{det}}} \rho_{\mathrm{L}}^2 \quad \text{with} \quad \rho_{\mathrm{L}}^2 \equiv \mathcal{A}_Y^\mu \mathcal{M}_{\mu\nu}^Y \mathcal{A}_Y^\nu , \quad (70)$$

where we refer to the (single-IFO) SNR ρ_{L} as the "line SNR".

The signal candidates are generated for a fixed SNR of $\rho_{\mathrm{S}} = 6$, and a data length of $T = 25\,\mathrm{h}$ is assumed. This signal strength is chosen to be representative of reasonably detectable signals in a wide-parameter-space search. In such a search we would require a low (single-trial) false-alarm threshold p_{FA} in order to consider a candidate as significant. The choice of $\rho_{\mathrm{S}} = 6$ corresponds to a detection probability of $p_{\mathrm{det}} \approx 70\%$ at a false-alarm probability of $p_{\mathrm{FA}} = 10^{-6}$ in Gaussian noise (for example, see Fig. 2).

The signal amplitude parameters are drawn uniformly in $\cos\iota \in [-1, 1]$, $\psi \in [-\pi/4, \pi/4]$ and $\phi_0 \in [0, 2\pi]$. The sky position is drawn isotropically over the sky, and $(h_0/\sqrt{S_{\mathrm{n}}})$ is determined by the fixed signal SNR of $\rho_{\mathrm{S}} = 6$ according to Eq. (69). The line draws use the same prior distributions, but the signal is added to only one detector, and $(h_0/\sqrt{S_{\mathrm{n}}})$ is determined by fixing a (single-IFO) line SNR ρ_{L} according to Eq. (70).

In each simulation we generate 10^7 noise candidates and 10^7 noise+signal candidates for two detectors, LIGO H1and L1. These detectors are assumed here to have identical sensitivity. Lines are only injected into H1without loss of generality. We consider three examples of noise populations:

- pure Gaussian noise without lines ($f_{\mathrm{L}} = 0$, $\rho_{\mathrm{L}} = 0$)

- 10% line contamination in H1 ($f_{\mathrm{L}}^{\mathrm{H1}} = 0.1$, $f_{\mathrm{L}}^{\mathrm{L1}} = 0$) with line SNR of $\rho_{\mathrm{L}} = 9$,

- 10% line contamination in H1 ($f_{\mathrm{L}}^{\mathrm{H1}} = 0.1$, $f_{\mathrm{L}}^{\mathrm{L1}} = 0$) with line SNR of $\rho_{\mathrm{L}} = 15$.

The line SNR of $\rho_{\mathrm{L}} = 9$ corresponds to lines that are marginally stronger than the injected signals, namely $E[2\mathcal{F}]_{\mathcal{H}_{\mathrm{L}}} \approx 44.5$ from Eq. (70)), while $E[2\mathcal{F}]_{\mathcal{H}_{\mathrm{S}}} = 40$ from Eq. (A7)). The lines with $\rho_{\mathrm{L}} = 15$ are substantially stronger (namely $E[2\mathcal{F}]_{\mathcal{H}_{\mathrm{L}}} \approx 117$) than the injected signals.

Note that for the synthesized statistics we cannot use the line-prior estimation method for o_{LG}^X of Sec. VI A. Instead we assume "perfect tuning": in the Gaussian-noise example we set $o_{\mathrm{LG}}^X = 0.001$ for $X = \mathrm{H1}, \mathrm{L1}$, and in the two line examples we use $p_{\mathrm{L}}^{\mathrm{H1}} = f_{\mathrm{L}}^{\mathrm{H1}} = 0.1$ (therefore $o_{\mathrm{LG}}^{\mathrm{H1}} = 1/9$) and $o_{\mathrm{LG}}^{\mathrm{L1}} = 0.001$ (no lines were injected into L1).
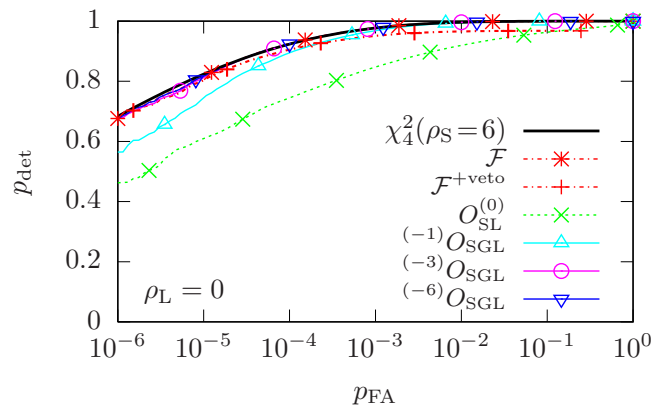

FIG. 2. Detection probability p_{det} as a function of false-alarm p_{FA} of different synthesized statistics, for a signal population of fixed SNR of $\rho_{\mathrm{S}} = 6$ in pure Gaussian noise ($f_{\mathrm{L}} = 0$, $\rho_{\mathrm{L}} = 0$). Statistical errors are similar to the line width.

In Gaussian noise the coherent $\mathcal{F}$-statistic is close to optimal [4, 6], and follows a χ^2 distribution with 4 degrees of freedom and non-centrality parameter ρ_{S}^2, which we denote as $\chi_4^2(\rho_{\mathrm{S}})$. This is plotted as a thick solid line in Figs. 2 and 3 for the signal population of $\rho_{\mathrm{S}} = 6$.

In the Gaussian-noise example shown in Fig. 2, the $\mathcal{F}$-statistic follows closely the theoretical prediction, while the (untuned) line-veto statistic $O_{\mathrm{SL}}^{(0)}$ is notably less powerful. The line-robust statistics $^{(-n)}O_{\mathrm{SGL}}$ increasingly approach the $\mathcal{F}$-statistic performance with decreasing
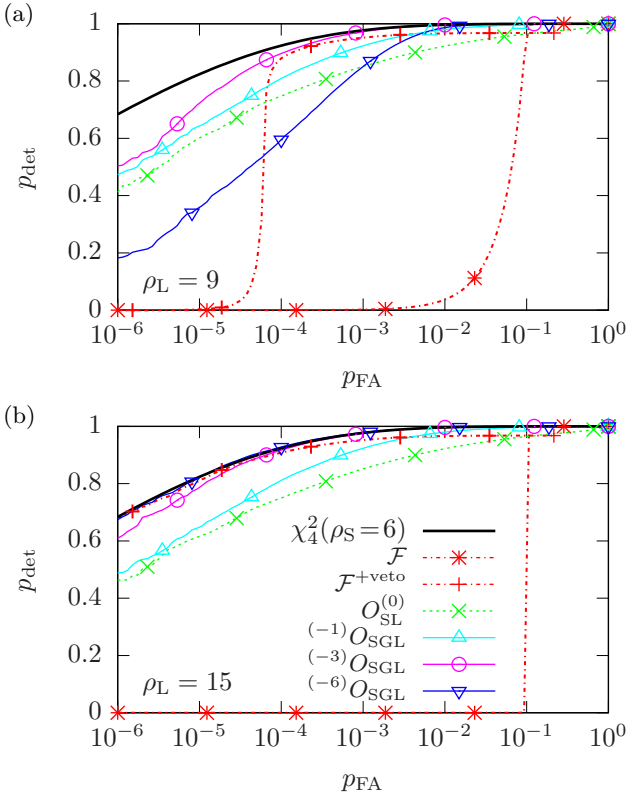
FIG. 3. Detection probability p_{det} as a function of false-alarm p_{FA} for different synthesized statistics, for a signal population with fixed SNR of $\rho_{\mathrm{S}} = 6$ in Gaussian noise with 10% line contamination, with line-SNR of (a) $\rho_{\mathrm{L}} = 9$ and (b) $\rho_{\mathrm{L}} = 15$. Statistical errors are similar to the line width.

FIG. 4. Comparison of "tuned" statistics $^{(-n)}O_{\mathrm{SGL}}$ (*solid lines*) using "perfect knowledge" line-priors $\{o_{\mathrm{LG}}^{\mathrm{H1}} = 1/9, o_{\mathrm{LG}}^{\mathrm{L1}} = 10^{-3}\}$ versus "untuned" statistic $^{(-n)}O_{\mathrm{SGL}}^{(0)}$ (*dashed lines*) using uninformative line-priors $o_{\mathrm{LG}}^{X} = 1$. Detection probability p_{det} as a function of false-alarm p_{FA} of different synthesized statistics, for a signal population with fixed SNR of $\rho_{\mathrm{S}} = 6$ in Gaussian noise with 10% line contamination, with line-SNR of (a) $\rho_{\mathrm{L}} = 9$ and (b) $\rho_{\mathrm{L}} = 15$. Statistical errors are similar to the line width.

$p_{\mathrm{FA}*}^{(0)}$, i.e., increasing transition scale $\mathcal{F}_{*}^{(0)}$. In particular, starting from $^{(-3)}O_{\mathrm{SGL}}$ (corresponding to a transition scale of $\mathcal{F}_{*}^{(0)} \approx 9.23$), there are no appreciable losses in detection probability p_{det} over the false-alarm range $p_{\mathrm{FA}} \in [10^{-6}, 1]$.

At low p_{FA}, the $\mathcal{F}^{+\mathrm{veto}}$-statistic performs almost optimally, while there are some losses above $p_{\mathrm{FA}} \gtrsim 10^{-4}$. These are due to $\mathcal{F}^{+\mathrm{veto}}$ containing intrinsic upper bounds on the achievable p_{FA} and p_{det} as a result of vetoing a finite fraction of candidates.

The performance in the two examples with 10% line contamination is shown in Fig. 3. Here the $\mathcal{F}$-statistic is found to perform substantially worse than in Gaussian noise at false-alarm probabilities below $p_{\mathrm{FA}} \lesssim 0.1$. This is due to the fact that in 10% of the noise cases the false-alarm threshold is set by the line population, which is either difficult (for $\rho_{\mathrm{L}} = 9$, left plot) or almost impossible (for $\rho_{\mathrm{L}} = 15$, right plot) for the $\mathcal{F}$-statistic to cross for signals with SNR of $\rho_{\mathrm{S}} = 6$.

We observe that the $\mathcal{F}^{+\mathrm{veto}}$-statistic starts to fail below false-alarm levels of $p_{\mathrm{FA}} \lesssim 10^{-4}$ in the case of weaker lines with $\rho_{\mathrm{L}} = 9$ (see Fig. 3(a)). This can be understood as follows: For the $\rho_{\mathrm{L}} = 9$ line population, we find that a fraction of $\sim 6 \times 10^{-4}$ of line candidates survive the veto.
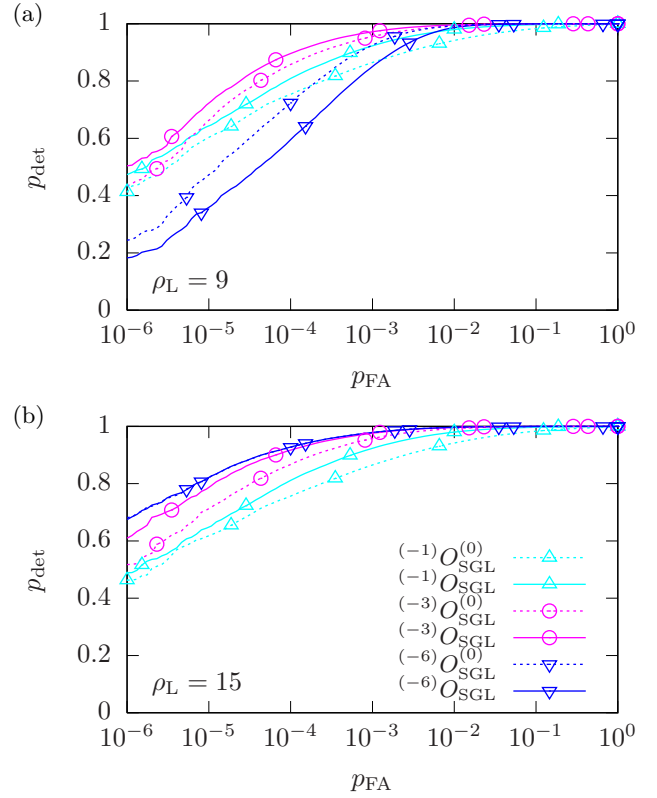
Given that lines are present in 10% of the noise cases, this means that a fraction of $\sim 6 \times 10^{-5}$ of noise candidates after the veto still originate from lines, whose high $\mathcal{F}$-statistic values are hard to cross for signal candidates. Therefore, the detection probability drops towards zero at false-alarm probabilities below this level.

The same effect is also present for stronger lines, but the corresponding "failure" threshold is pushed to lower values. For example, for $\rho_{\mathrm{L}} = 12$ it would happen only below $p_{\mathrm{FA}} \lesssim 10^{-6}$, while for $\rho_{\mathrm{L}} = 15$ it is too low to be resolvable by 10^7 random draws.

The behavior of the line-robust statistics $^{(-n)}O_{\mathrm{SGL}}$ depends on the choice of transition scale. In the case of lines with $\rho_{\mathrm{L}} = 9$, shown in Fig. 3(a), the statistic $^{(-3)}O_{\mathrm{SGL}}$ performs best, while using either lower or higher values of $p_{\mathrm{FA}*}^{(0)}$ is less powerful at low false-alarm probabilities. In the case of stronger lines with $\rho_{\mathrm{L}} = 15$, shown in Fig. 3(b), the statistic $^{(-6)}O_{\mathrm{SGL}}$ performs almost optimally, with $^{(-3)}O_{\mathrm{SGL}}$ performing only slightly worse.

The line-veto statistic $O_{\mathrm{SL}}^{(0)}$ performs somewhat poorly in all three examples shown (Figs. 2 and 3). This is

not surprising, given that at most 10% of noise draws contain a line, while O_{SL} would only be optimal for a noise population consisting exclusively of lines.

Figure 4 shows the effect of "tuning" the prior line odds o_{LG}^X, using the same line populations as in Fig. 3. We see that the untuned statistics $O_{\mathrm{SL}}^{(0)}$ and $^{(-n)}O_{\mathrm{SGL}}^{(0)}$ using uninformative line-odds $o_{\mathrm{LG}}^X = 1$ perform reasonably well compared to O_{SL} and $^{(-n)}O_{\mathrm{SGL}}$, which are based on "perfect-knowledge" tuning. Note that tuning of o_{LG}^X can sometimes also decrease the detection power of a statistic, particularly in cases where the choice of the transition scale $\mathcal{F}_*^{(0)}(p_{\mathrm{FA}*}^{(0)})$ is a poor fit to the actual line population. This can be seen in the case of $^{(-6)}O_{\mathrm{SGL}}$ with lines of SNR $\rho_{\mathrm{L}} = 9$, as shown in Fig. 4(a). In cases where $\mathcal{F}_*^{(0)}(p_{\mathrm{FA}*}^{(0)})$ is a good match to the line population, the tuning of o_{LG}^X can yield gains in detection power of up to 5–10%.

B. Tests using LIGO S5 data

Here we conduct a study using LIGO S5 data sets as noise. We inject signals and search for them using methods similar to those of actual CW searches. Instead of ROC curves, i.e., $p_{\mathrm{det}}(p_{\mathrm{FA}})$, we present results in the form of detection efficiency as a function of signal strength, i.e., $p_{\mathrm{det}}(h_0/\sqrt{S_{\mathrm{n}}})$. This form is more suitable to assess improvements in sensitivity, which is typically expressed as the weakest signal h_0 detectable with a certain confidence p_{det}. The injection and detection procedure used here is modeled after those commonly employed for estimating upper limits in CW searches such as Refs. [11, 28].

Simulated CW signals are added to the data using the Makefakedata_v4 code [43]. The resulting data set is analyzed both coherently and semi-coherently using HierarchSearchGCT [43], a StackSlide implementation based on the "global correlations" method of Ref. [44]. We have extended this code to also compute the new statistics $\widehat{O}_{\mathrm{SL}}(\boldsymbol{x})$ and $\widehat{O}_{\mathrm{SGL}}(\boldsymbol{x})$, in addition to $\widehat{\mathcal{F}}(\boldsymbol{x})$. For the coherent search we use shorter subsets of the data, and the coherent statistics are simply obtained as the special case $N_{\mathrm{seg}} = 1$.

The tuning of the line priors $\widehat{o}_{\mathrm{LG}}^X$ in $\widehat{O}_{\mathrm{SL}}$ and $^{(-n)}\widehat{O}_{\mathrm{SGL}}$ is based on the method described in Sec. VI A, namely Eqs. (64), (65), and (66). As explained in Sec. VI B, we fix the transition scale $\widehat{\mathcal{F}}_*^{(0)}$ of $\widehat{O}_{\mathrm{SGL}}$ according to its performance in Gaussian noise. Specifically, we perform injections on simulated Gaussian noise and analyze them as described below for several values of $p_{\mathrm{FA}*}^{(0)}$. We then chose the highest $p_{\mathrm{FA}*}^{(0)}$ value such that the achieved performance is indistinguishable within statistical uncertainties from that of the $\mathcal{F}$-statistic. As a result of this we select $^{(-6)}O_{\mathrm{SGL}}$, with the false-alarm level of $p_{\mathrm{FA}*}^{(0)} = 10^{-6}$ corresponding to a transition scale of $\mathcal{F}_*^{(0)}(N_{\mathrm{seg}} = 1) \approx 16.7$ and $\widehat{\mathcal{F}}_*(N_{\mathrm{seg}} = 84) \approx 237.0$, respectively.

Data selection

We use four narrow frequency bands of LIGO S5 data. These bands are chosen depending on how severely they appear to be affected by lines:

(a) a "quiet" band where the distribution of the data is very close to Gaussian,

(b) a band with a single line in L1,

(c) a band with a single line in L1, narrower than in (b),

(d) a band with multiple disturbances in H1,

The normalized SFT power $\mathcal{P}^X(f)$ of Eq. (63) for each of the four bands is shown in Fig. 5 for the coherent case, and in Fig. 7 for the semi-coherent case. More details about these sample frequency bands are given in Table I and II, respectively.

The data sets are taken from the first year of the LIGO S5 science run. For the semi-coherent searches we use $N_{\mathrm{seg}} = 84$ data segments, spanning $T = 25\,\mathrm{h}$ each, while the coherent searches use only a single segment. These segments were originally selected for the Einstein@Home [45] search described in Ref. [11]. Since CW searches on this data have not found any signals [11, 27, 46], we consider it as a *pure noise* set for the purpose of this study.

Signal injection and detection criterion

The search setup used here is different from that of Ref. [11], and employs the HierarchSearchGCT code instead of the Hough-transform [38]. This code is used in recent and ongoing wide-parameter-space searches such as Refs. [28, 45].

The grid spacings in frequency and spindown are $\delta f \approx 1.6 \times 10^{-6}\,\mathrm{Hz}$ and $\delta\dot{f} \approx 5.8 \times 10^{-11}\,\mathrm{Hz}/s$, respectively. The angular sky-grid spacings are approximately $0.15\,\mathrm{rad}$ at $f = 54\,\mathrm{Hz}$, and scale with frequency as $1/f$.

We find that this template bank yields an average relative loss of SNR^2 (also known as mismatch) of $m \sim 0.6$ in the semi-coherent searches and of $m \lesssim 0.05$ in the coherent searches.

We first perform searches on the data without any injections, covering the whole sky in each of the four frequency bands of width $\Delta f = 50\,\mathrm{mHz}$ (see f_{inj} in Tables I and II), and a fixed band $[-\Delta\dot{f}, 0]$ in spindown $\dot{f}$, with $\Delta\dot{f} \approx 2.6 \times 10^{-9}\,\mathrm{Hz}/s$.

For each of the four statistics $\{\widehat{\mathcal{F}}, \widehat{\mathcal{F}}^{+\mathrm{veto}}, \widehat{O}_{\mathrm{SL}}, \widehat{O}_{\mathrm{SGL}}\}$ we record the loudest noise candidate over the whole template grid. A signal will be considered as detected with a given statistic if its highest value exceeds this noise value. This definition of detection is equivalent to the common method of setting loudest-event upper limits, employed for example in Ref. [11].

Label	$f_{\rm inj}$ [Hz]	$f_{\rm SFT}$ [Hz]	$t_{\rm start}$ [s]	max $2\mathcal{F}_{\rm noise}$	Detector X	$N_{\rm SFT}^X$	$\sqrt{S_{\rm n}^X}$ [Hz$^{-1/2}$]	$\mathcal{P}_{\rm thr}^X$	$o_{\rm LG}^X$
$(\widetilde{\rm a})$	$[54.20, 54.25]$	$[54.19, 54.26]$	835120582	38.04	H1	47	2.51×10^{-22}	1.875	0.001
					L1	39	1.48×10^{-22}	1.960	0.001
$(\widetilde{\rm b})$	$[66.50, 66.55]$	$[66.49, 66.56]$	844876223	81.68	H1	47	1.19×10^{-22}	1.875	0.001
					L1	36	1.41×10^{-22}	2.000	0.073
$(\widetilde{\rm c})$	$[69.70, 69.75]$	$[69.69, 69.76]$	821912087	122.54	H1	46	1.35×10^{-22}	1.884	0.001
					L1	40	1.50×10^{-22}	1.948	0.015
$(\widetilde{\rm d})$	$[58.50, 58.55]$	$[58.49, 58.56]$	827366996	104.46	H1	48	2.58×10^{-22}	1.866	0.585
					L1	40	2.40×10^{-22}	1.948	0.001

TABLE I. Data used for tests of the coherent statistics in Sec. VII B. All data is taken from the first year of the LIGO S5 run. CW signals are injected with frequencies $f \in f_{\rm inj}$, while $f_{\rm SFT}$ denotes the SFT frequency range used for the search and the prior line estimation. Each data set starts at a GPS time of $t_{\rm start}$ and spans 25 hours, containing $N_{\rm SFT}^X$ SFTs of duration $T_{\rm SFT} = 1800\,{\rm s}$ from each detector. The column labeled max $2\mathcal{F}_{\rm noise}$ shows the corresponding highest multi-detector $2\mathcal{F}$ value without injections. The average noise PSD (harmonic mean over SFTs, arithmetic mean over frequency bins) is $S_{\rm n}^X$. The multi-detector PSD $S_{\rm n}$ is obtained as the harmonic mean of $S_{\rm n}^X$ over X and is not tabulated. The column $\mathcal{P}_{\rm thr}^X$ gives the threshold on the normalized SFT power $\mathcal{P}^X$ at $p_{\rm FA,\mathcal{P}} = 10^{-9}$, which is used to estimate the prior line-odds $o_{\rm LG}^X$ as described in Sec. VI A.

Label	$f_{\rm inj}$ [Hz]	$f_{\rm SFT}$ [Hz]	max $2\overline{\mathcal{F}}_{\rm noise}$	Detector X	$N_{\rm SFT}^X$	$\sqrt{S_{\rm n}^X}$ [Hz$^{-1/2}$]	$\mathcal{P}_{\rm thr}^X$	$\widehat{o}_{\rm LG}^X$
$(\widehat{\rm a})$	$[54.20, 54.25]$	$[54.15, 54.30]$	6.51	H1	3781	2.54×10^{-22}	1.098	0.001
				L1	3456	1.81×10^{-22}	1.102	0.001
$(\widehat{\rm b})$	$[66.50, 66.55]$	$[66.44, 66.61]$	10.83	H1	3781	1.35×10^{-22}	1.098	0.001
				L1	3456	1.00×10^{-22}	1.102	0.047
$(\widehat{\rm c})$	$[69.70, 69.75]$	$[69.64, 69.81]$	83.48	H1	3781	1.15×10^{-22}	1.098	0.001
				L1	3456	9.08×10^{-23}	1.102	0.017
$(\widehat{\rm d})$	$[58.50, 58.55]$	$[58.45, 58.60]$	8.35	H1	3781	2.20×10^{-22}	1.098	1.743
				L1	3456	2.05×10^{-22}	1.102	0.001

TABLE II. Data used for tests of the semi-coherent statistics in Sec. VII B. All data is taken from the first year of the LIGO S5 run, corresponding to the segment selection used in an Einstein@Home search (S5R3) [11], spanning 381.04 days starting from GPS epoch $t_{\rm start} = 818845553$, containing $N_{\rm seg} = 84$ segments, each 25 hours long. The column labeled max $2\overline{\mathcal{F}}_{\rm noise}$ refers to the highest average multi-detector $2\overline{\mathcal{F}}$ value without injections (the average is over segments). The remaining labels are identical to those in Table I.

The signals are injected using the `Makefakedata_v4` code, with signal parameters randomly drawn from uniform distributions in the sky coordinates $\{\alpha, \delta\}$, inclination $\cos \iota$ and polarization angle ψ, and at varying signal amplitude h_0. The signal frequency and spindown are drawn uniformly from the bands used in the noise search. For each value of h_0 we perform 1000 injections. For each injection we search a small parameter-space volume containing the signal. This search region consists of a frequency band of $\Delta f = 1\,{\rm mHz}$, a spindown band of $\Delta \dot{f} \approx 2.3 \times 10^{-10}$ Hz/s and the 10 sky-grid points closest (in the metric sense [33]) to the injection.

Note that in (b) and (c) some of these injection searches do not use any data containing the narrow disturbances. Hence, the statements in this section apply to *bands that contain disturbances*, and not only to *sets of disturbed candidates*.

Results for coherent statistics

Figure 6 shows the detection efficiency $p_{\rm det}$ as a function of the scaled signal amplitude $h_0/\sqrt{S_{\rm n}}$, for the single-segment coherent statistics.

In the quiet band, shown in Fig. 6 $(\widetilde{\rm a})$, we find that the line-veto statistic $O_{\rm SL}^{(0)}$ has less detection power than the $\mathcal{F}$-statistic, as would be expected since it does not match the noise population. The conventional $\mathcal{F}^{+\rm veto}$-statistic is safer than $O_{\rm SL}^{(0)}$ and performs just as well as the pure $\mathcal{F}$-statistic. The line-robust statistic $O_{\rm SGL}$ performs equally well as $\mathcal{F}$ and $\mathcal{F}^{+\rm veto}$ on this line-free data set.

In the disturbed bands shown in Fig. 6 $(\widetilde{\rm b})$-$(\widetilde{\rm d})$, all statistics lose detection power to varying degrees. We find that the $\mathcal{F}^{+\rm veto}$-statistic is often able to recover most of the losses of the pure $\mathcal{F}$-statistic. The line-veto statistic $O_{\rm SL}^{(0)}$ performs similarly in case $(\widetilde{\rm b})$ and yields an improvement over $\mathcal{F}^{+\rm veto}$ in cases $(\widetilde{\rm c})$ and $(\widetilde{\rm d})$. However,
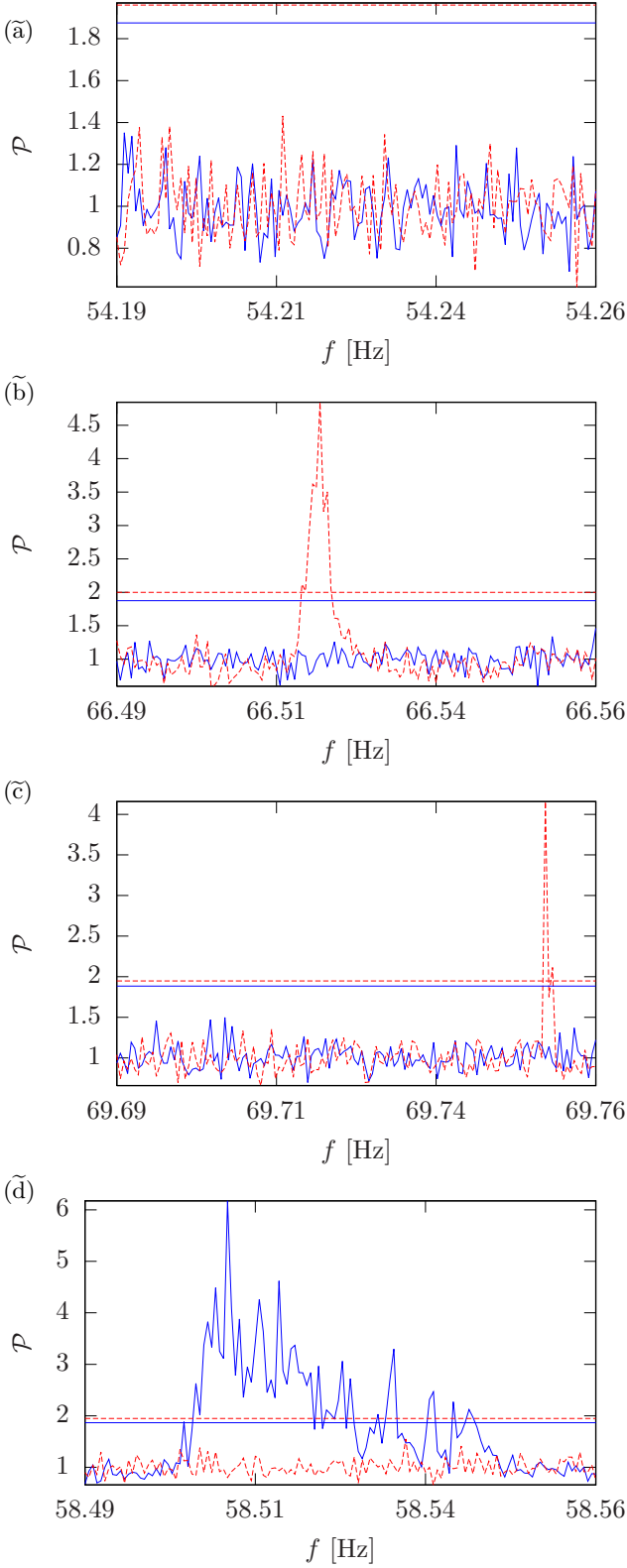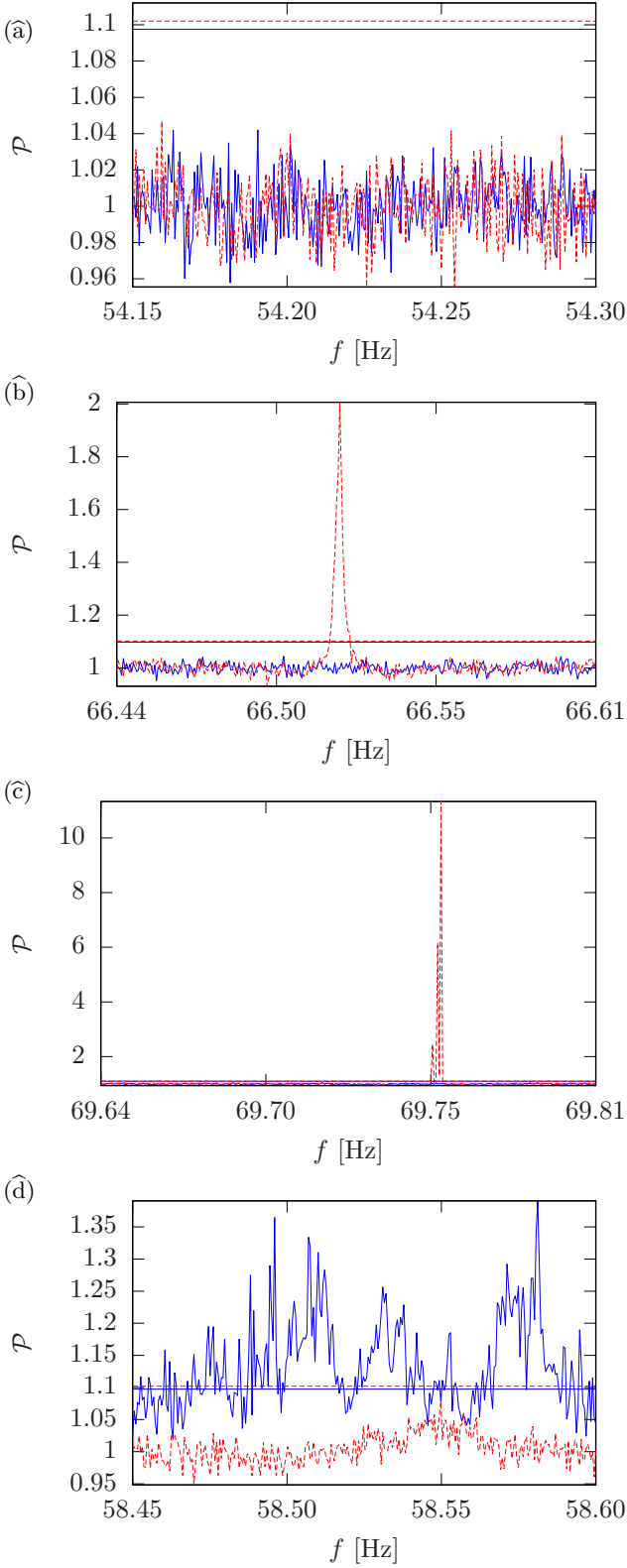
FIG. 5. Normalized average SFT power $\mathcal{P}^X(f)$ of Eq. (63) as a function of frequency f for LIGO H1 (*solid*) and L1 (*dashed*) data used in the coherent searches. The horizontal lines mark, for each detector, the threshold $\mathcal{P}^X_{\text{thr}}$ at $p_{\text{FA},\mathcal{P}} = 10^{-9}$ used in the line prior estimation. The panels show: ($\widetilde{\text{a}}$) a quiet band, ($\widetilde{\text{b}}$), ($\widetilde{\text{c}}$) two bands with lines, ($\widetilde{\text{d}}$) a band with multiple disturbances. See Table I for more details on these data sets.
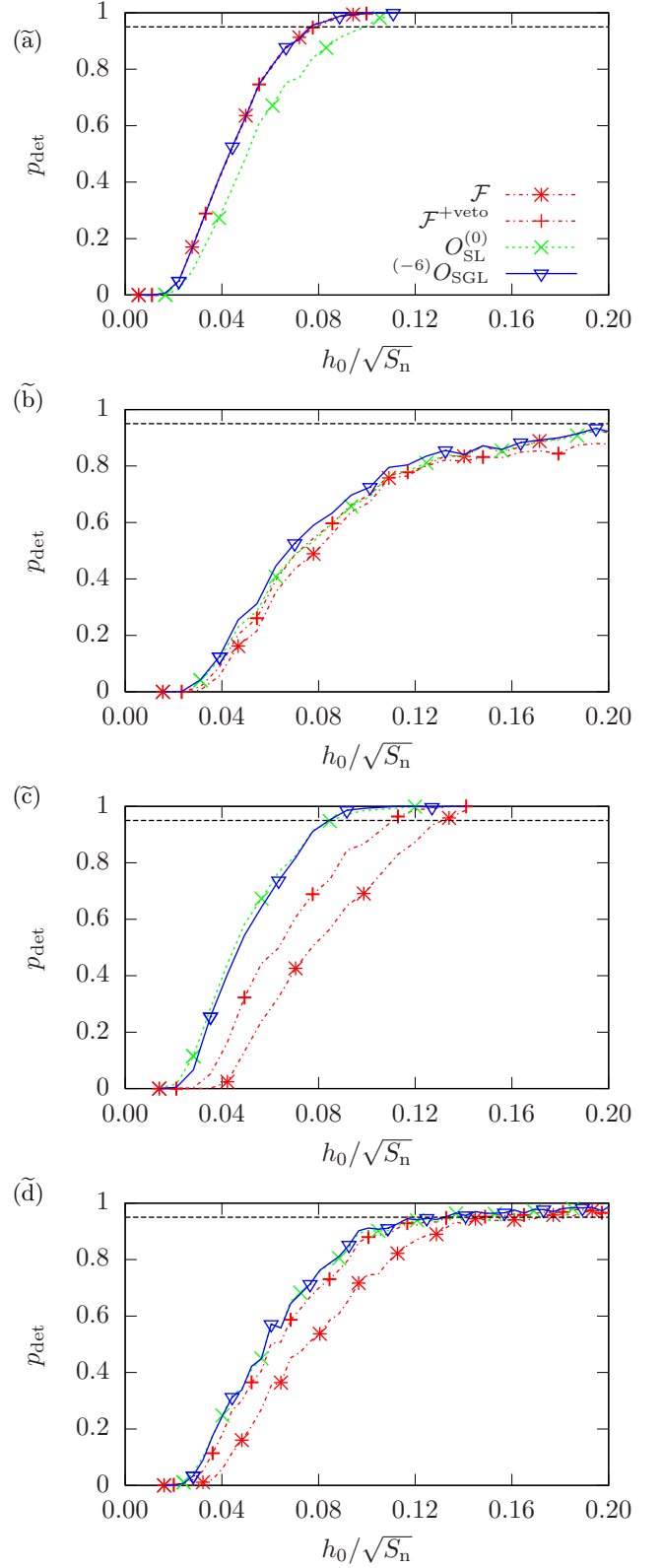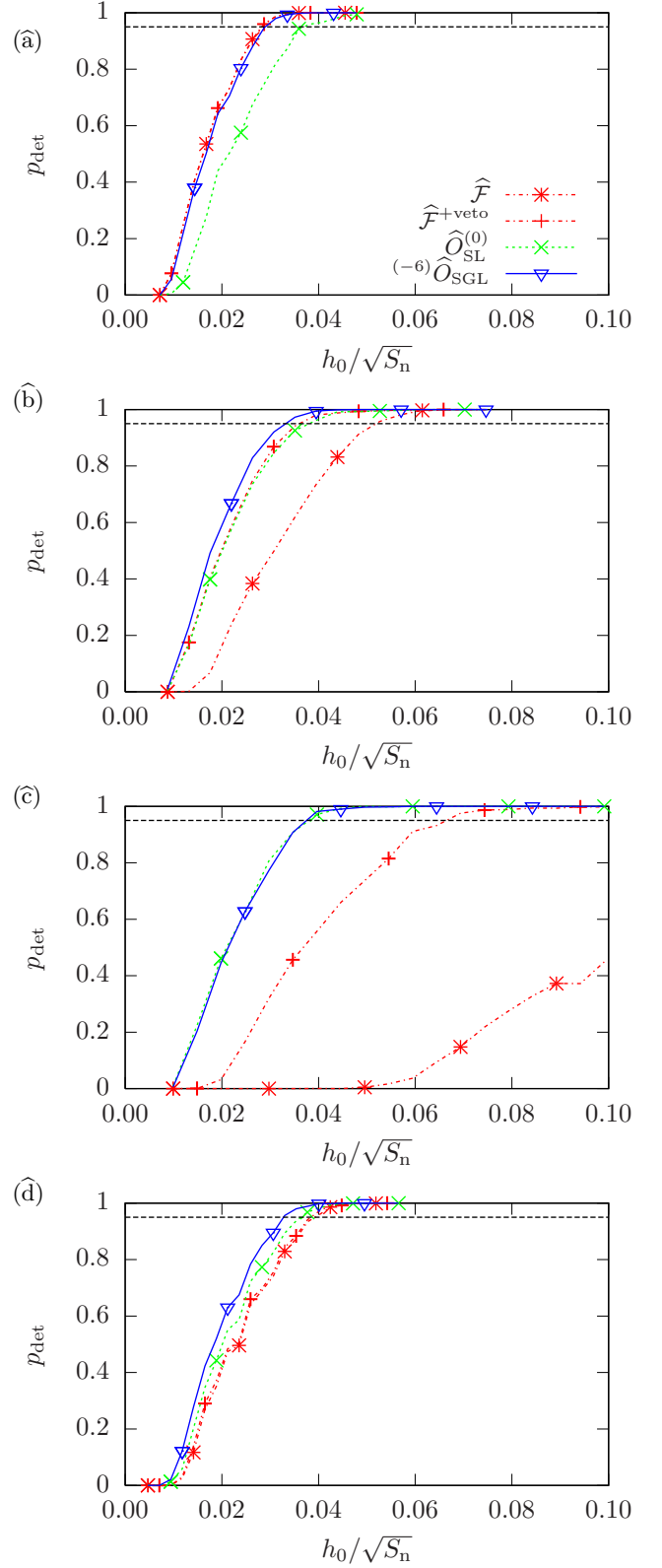
FIG. 6. Detection efficiency p_{det} as a function of scaled signal amplitude $h_0/\sqrt{S_{\text{n}}}$ for four different coherent statistics: $\mathcal{F}$, $\mathcal{F}^{+\text{veto}}$, $O^{(0)}_{\text{SL}}$, and $^{(-6)}O_{\text{SGL}}$. Statistical errors are similar to the size of the symbols. The dashed horizontal line marks the 95% detection probability level. The panels show: ($\widetilde{\text{a}}$) a quiet band, ($\widetilde{\text{b}}$), ($\widetilde{\text{c}}$) two bands with lines, ($\widetilde{\text{d}}$) a band with multiple disturbances. See Fig. 5 and Table I for more details on these data sets.

FIG. 7. Normalized average SFT power $\mathcal{P}^X(f)$ of Eq. (63) as a function of frequency f for LIGO H1 (*solid*) and L1 (*dashed*) data used in the semi-coherent searches. The horizontal lines mark, for each detector, the threshold $\mathcal{P}^X_{\text{thr}}$ at $p_{\text{FA},\mathcal{P}} = 10^{-9}$ used in the line prior estimation. The panels show: (â) a quiet band, (b̂), (ĉ) two bands with lines, (d̂) a band with multiple disturbances. See Table II for more details on these data sets.

FIG. 8. Detection efficiency p_{det} as a function of scaled signal amplitude $h_0/\sqrt{S_{\text{n}}}$ for four different semi-coherent statistics: $\widehat{\mathcal{F}}$, $\widehat{\mathcal{F}}^{+\text{veto}}$, $\widehat{O}^{(0)}_{\text{SL}}$, and $^{(-6)}\widehat{O}_{\text{SGL}}$. Statistical errors are similar to the size of the symbols. The dashed horizontal line marks the 95% detection probability level. The panels show: (â) a quiet band, (b̂), (ĉ) two bands with lines, (d̂) a band with multiple disturbances. See Fig. 7 and Table II for more details on these data sets.

these cases show O_{SGL} to be more robust than either of the simpler vetoes.

Summarizing these results, we see that the line-robust statistic O_{SGL} consistently shows the best performance over the different types of data: it is more robust to varying kinds of disturbances than $\mathcal{F}^{+\mathrm{veto}}$ and safer in Gaussian noise than $O_{\mathrm{SL}}^{(0)}$.

Results for semi-coherent statistics

Figure 8 shows the detection efficiency p_{det} as a function of $h_0/\sqrt{S_{\mathrm{n}}}$ for the semi-coherent statistics over the full data set. Qualitatively, we find very similar results to the coherent case of Fig. 6.

For the quiet band, shown in Fig. 8 (â), we find that the simple line-veto $\widehat{O}_{\mathrm{SL}}^{(0)}$ loses a significant fraction of detection power compared to the semi-coherent $\widehat{\mathcal{F}}$-statistic and to $\widehat{\mathcal{F}}^{+\mathrm{veto}}$, while the line-robust statistic $\widehat{O}_{\mathrm{SGL}}$ does not show any significant degradation.

In the bands with noise disturbances (Fig. 6 (b̂)-(d̂)), it is again the $\widehat{\mathcal{F}}$-statistic which suffers the most. These examples show the line-robust statistic $\widehat{O}_{\mathrm{SGL}}$ consistently performing better than $\widehat{\mathcal{F}}$ and as well as or better than either $\widehat{O}_{\mathrm{SL}}^{(0)}$ or $\widehat{\mathcal{F}}^{+\mathrm{veto}}$ in all the disturbed bands. The largest improvement is found in the example shown in Fig. 8 (ĉ), where the signal amplitude at $95\,\%$ detection probability is nearly two times smaller for $\widehat{O}_{\mathrm{SGL}}$ compared to $\widehat{\mathcal{F}}^{+\mathrm{veto}}$.

VIII.   CONCLUSIONS

We have extended the standard derivation of the $\mathcal{F}$-statistic by adding an explicit simple line hypothesis to the standard Gaussian-noise hypothesis, namely a CW-signal-like disturbance in a single detector. More work would be required to deal with coincident disturbances in multiple detectors.

Using the Bayesian framework we have derived two new detection statistics: a "line-veto" statistic O_{SL}, which complements the $\mathcal{F}$-statistic and may be appropriate for the follow-up of strong outliers, and a new line-robust detection statistic O_{SGL}, which contains both $\mathcal{F}$ and O_{SL} as limiting cases. We have also generalized both statistics to semi-coherent searches.

The line-robust O_{SGL} requires choosing several prior parameters. We have found in particular that the performance of O_{SGL} is sensitive to $\mathcal{F}_*^{(0)}$, which regulates the transition scale between $\mathcal{F}$ and O_{SL}. This parameter stems from a rather unphysical prior in the $\mathcal{F}$-statistic derivation [6], and we could therefore only provide an *ad-hoc* empirical prescription for choosing it. Further work to improve on this prior could also result in increased robustness when the detectors are not equally sensitive.

The remaining parameters are more straightforward to interpret, as they encode the prior probability of line artifacts. For these we have tested both an ignorance prior and a simple adaptive tuning method.

We have tested the detection power of the new statistics on synthetic candidates, where both signal and noise match our hypotheses, and on simulated signals injected into LIGO S5 data. In both cases we have found that, with a reasonable choice of transition scale, O_{SGL} is consistently the most robust in the presence of various types of instrumental artifacts. In particular, it consistently equals or surpasses the performance of the popular *ad-hoc* $\mathcal{F}$-statistic consistency veto.

Combined with its close-to-optimal performance in undisturbed data, this makes O_{SGL} a promising statistic for analyzing broad-band, diverse data sets.

Appendix A: Expectation value of the $\mathcal{F}$-statistic under the line hypothesis $\mathcal{H}_{\mathrm{L}}$

In order to derive the expectation value of the $\mathcal{F}$-statistic under the hypothesis $\mathcal{H}_{\mathrm{L}}$ given in Eq. (19), we consider the more general case of a signal with detector-dependent amplitude parameters $\mathcal{A}_X^\mu$. The signal case corresponds to $\mathcal{A}_X^\mu = \mathcal{A}^\mu$ for all X, while the line hypothesis corresponds to $\mathcal{A}_X^\mu = \mathcal{A}_Y^\mu \delta_{XY}$, with Y denoting the detector containing the instrumental line. By using the factorization of Eq. (6), we can express the data in the general case as

$$x^X = n^X + \mathcal{A}_X^\alpha\, h_\alpha^X\,, \tag{A1}$$

with *no* automatic summation over repeated detector indices $X, X', \dots$. Hence, the projections of Eq. (9) for detector X read as

$$x_\mu^X \equiv \left(x^X | h_\mu^X\right) = n_\mu^X + \mathcal{A}_X^\alpha \mathcal{M}_{\alpha\mu}^X\,, \tag{A2}$$

where we define $n_\mu^X \equiv \left(n^X | h_\mu^X\right)$, and from Eq. (3) we have

$$x_\mu = \sum_X x_\mu^X\,, \quad \text{and} \quad \mathcal{M}_{\mu\nu} = \sum_X \mathcal{M}_{\mu\nu}^X\,. \tag{A3}$$

The expectation value of the $\mathcal{F}$-statistic of Eq. (13) can therefore be written as

$$E\left[2\mathcal{F}\right] = \mathcal{M}^{\mu\nu}\, E\left[x_\mu x_\nu\right]\,, \qquad (A4)$$

where

$$E\left[x_\mu x_\nu\right] = \sum_{XX'} E\left[x_\mu^X\, x_\nu^{X'}\right]$$
$$= \mathcal{M}_{\mu\nu} + \sum_{XX'} \mathcal{A}_X^\alpha \mathcal{M}_{\alpha\mu}^X \mathcal{M}_{\nu\beta}^{X'} \mathcal{A}_{X'}^\beta\,, \qquad (A5)$$

using the noise expectation values $E\left[n_\mu^X n_\nu^{X'}\right] = \mathcal{M}_{\mu\nu}^X \delta^{XX'}$ for uncorrelated noise between detectors, and $E\left[n_\mu^X\right] = 0$, assuming zero-mean noise n^X.

Given that $\mathcal{M}^{\mu\nu}\mathcal{M}_{\mu\nu} = 4$, we obtain the general result for the expectation value of the $\mathcal{F}$-statistic:

$$E\left[2\mathcal{F}\right] = 4 + \mathcal{M}^{\mu\nu} \sum_{XX'} \mathcal{A}_X^\alpha \mathcal{M}_{\alpha\mu}^X \mathcal{M}_{\nu\beta}^{X'} \mathcal{A}_Y^\beta\,. \qquad (A6)$$

A CW signal has consistent amplitudes in all detectors, i.e., $\mathcal{A}_X^\alpha = \mathcal{A}^\alpha$, and this leads to the well-known result

for $E\left[2\mathcal{F}\right]$ in terms of the signal SNR ρ_S:

$$E\left[2\mathcal{F}\right]_{\mathcal{H}_S} = 4 + \rho_S^2\,, \quad \text{with} \quad \rho_S^2 = \mathcal{A}^\mu \mathcal{M}_{\mu\nu} \mathcal{A}^\nu\,. \qquad (A7)$$

Under the line hypothesis $\mathcal{H}_L$, i.e., $\mathcal{A}_X^\mu = \mathcal{A}_Y^\mu \delta_{XY}$, we find

$$E\left[2\mathcal{F}\right]_{\mathcal{H}_L} = 4 + \mathcal{M}^{\mu\nu} \mathcal{A}_Y^\alpha \mathcal{M}_{\alpha\mu}^Y \mathcal{M}_{\nu\beta}^Y \mathcal{A}_Y^\beta\,. \qquad (A8)$$

In the special case of all detectors having identical antenna-pattern matrices, i.e., $\mathcal{M}_{\mu\nu}^X = \mathcal{M}_{\mu\nu}^1$, we have $\mathcal{M}_{\mu\nu} = N_{\det}\,\mathcal{M}_{\mu\nu}^1$, and therefore $\mathcal{M}^{\mu\nu} = \frac{1}{N_{\det}}\mathcal{M}_1^{\mu\nu}$. This results in

$$E\left[2\mathcal{F}\right]_{\mathcal{H}_L} = 4 + \frac{1}{N_{\det}}\,\rho_L^2\,, \qquad (A9)$$

with

$$\rho_L^2 \equiv \mathcal{A}_Y^\mu \mathcal{M}_{\mu\nu}^Y \mathcal{A}_Y^\nu\,, \qquad (A10)$$

defining the (single-IFO) "line SNR" in detector Y. The scaling of Eq. (A9) should still approximately hold for detectors with not-too-different antenna-pattern matrices $\mathcal{M}_{\mu\nu}^X$.

This result shows that a CW-like disturbance with SNR ρ_L in a single detector is not completely suppressed in the multi-detector $\mathcal{F}$-statistic, but is only reduced to an effective multi-detector SNR of approximately $\rho_L/\sqrt{N_{\det}}$.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration), Rep. Prog. Phys. **72**, 076901 (2009), arXiv:0711.3041 [gr-qc].
[2] T. Accadia *et al.*, Class. Quant. Grav. **28**, 114002 (2011).
[3] H. Grote (for the LIGO Scientific Collaboration), Class. Quant. Grav. **27**, 084003 (2010).
[4] P. Jaranowski, A. Królak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998), arXiv:gr-qc/9804014.
[5] C. Cutler and B. F. Schutz, Phys. Rev. D **72**, 063006 (2005), arXiv:gr-qc/0504011.
[6] R. Prix and B. Krishnan, Class. Quant. Grav. **26**, 204013 (2009), arXiv:0907.2569 [gr-qc].
[7] L. Blackburn *et al.*, Class. Quant. Grav. **25**, 184004 (2008), arXiv:0804.0800 [gr-qc].
[8] J. Aasi *et al.* (The LIGO Scientific Collaboration and The Virgo Collaboration), Class. Quant. Grav. **29**, 155002 (2012), arXiv:1203.5613 [gr-qc].
[9] J. Slutsky *et al.*, Class. Quant. Grav. **27**, 165023 (2010), arXiv:1004.0998 [gr-qc].
[10] T. Prestegard, E. Thrane, N. L. Christensen, M. W. Coughlin, B. Hubbert, S. Kandhasamy, E. MacAyeal, and V. Mandic, Class. Quant. Grav. **29**, 095018 (2012), arXiv:1111.1631 [astro-ph.IM].
[11] J. Aasi *et al.* (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. D **87**, 042001 (2013), arXiv:1207.7176 [gr-qc].
[12] T. Accadia *et al.*, J. Phys. Conf. Ser. **363**, 012037 (2012).
[13] K. C. Cannon, Class. Quant. Grav. **25**, 105024 (2008).
[14] B. Allen, Phys. Rev. D **71**, 062001 (2005), arXiv:gr-qc/0405045.
[15] L. Wen and B. F. Schutz, Class. Quant. Grav. **22**, 1321 (2005), arXiv:gr-qc/0508042.
[16] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **72**, 082001 (2005), arXiv:gr-qc/0505041.
[17] S. Babak *et al.*, Phys. Rev. D **87**, 024033 (2013), arXiv:1208.3491 [gr-qc].
[18] I. W. Harry and S. Fairhurst, Phys. Rev. D **83**, 084002 (2011), arXiv:1012.4939 [gr-qc].
[19] J. Abadie *et al.* (LIGO Collaboration, Virgo Collaboration), Phys. Rev. D **85**, 082002 (2012), arXiv:1111.7314 [gr-qc].
[20] P. J. Sutton, G. Jones, S. Chatterji, P. Kalmus, I. Leonor, S. Poprocki, J. Rollins, A. Searle, L. Stein, M. Tinto, and M. Was, New Journal of Physics **12**, 053034 (2010), arXiv:0908.3665 [gr-qc].
[21] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **80**, 102001 (2009), arXiv:0905.0020 [gr-qc].
[22] J. Abadie *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. D **85**, 122007 (2012), arXiv:1202.2788 [gr-qc].
[23] J. Clark, I. S. Heng, M. Pitkin, and G. Woan, Phys. Rev. D **76**, 043003 (2007), arXiv:gr-qc/0703138.
[24] T. Dal Canton, S. Bhagwat, S. V. Dhurandhar, and A. Lundgren, Class. Quant. Grav. **31**, 015016 (2014), arXiv:1304.0008 [gr-qc].
[25] T. B. Littenberg and N. J. Cornish, Phys. Rev. D **82**, 103007 (2010), arXiv:1008.1577 [gr-qc].

[26] J. Veitch and A. Vecchio, Phys. Rev. D **81**, 062003 (2010), arXiv:0911.3820 [astro-ph.CO].

[27] J. Abadie *et al.* (The LIGO Scientific Collaboration and The Virgo Collaboration), Phys. Rev. D **85**, 022001 (2012), arXiv:1110.0208 [gr-qc].

[28] J. Aasi *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. D **88**, 102002 (2013), arXiv:1309.6221 [gr-qc].

[29] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **77**, 022001 (2008), arXiv:0708.3818 [gr-qc].

[30] J. Abadie *et al.*, Astrophys. J. **722**, 1504 (2010), arXiv:1006.2535 [gr-qc].

[31] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **79**, 022001 (2009), arXiv:0804.1747 [gr-qc].

[32] B. Behnke, *A Directed Search for Continuous Gravitational Waves from Unknown Isolated Neutron Stars at the Galactic Center*, Ph.D. thesis, Hannover University (2013).

[33] R. Prix, Phys. Rev. D **75**, 023004 (2007), arXiv:gr-qc/0606088.

[34] R. Prix for the LIGO Scientific Collaboration, in *Neutron Stars and Pulsars*, Astrophysics and Space Science Library, Vol. 357, edited by W. Becker (Springer Berlin Heidelberg, 2009) Chap. 24, p. 651, (LIGO-P060039-v3).

[35] R. Prix, S. Giampanis, and C. Messenger, Phys. Rev. D **84**, 023007 (2011), arXiv:1104.1704 [gr-qc].

[36] P. R. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000), arXiv:gr-qc/9812014.

[37] R. Prix and M. Shaltev, Phys. Rev. D **85**, 084010 (2012), arXiv:1201.4321 [gr-qc].

[38] B. Krishnan, A. M. Sintes, M. A. Papa, B. F. Schutz, S. Frasca, and C. Palomba, Phys. Rev. D **70**, 082001 (2004), arXiv:gr-qc/0407001.

[39] K. Wette, *Gravitational waves from accreting neutron stars and Cassiopeia A*, Ph.D. thesis, Department of Quantum Science, Research School of Physics and Engineering, Australian National University (2009).

[40] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **69**, 082004 (2004), arXiv:gr-qc/0308050 [gr-qc].

[41] E. T. Jaynes, *Probability Theory. The Logic of Science* (Cambridge University Press, 2003).

[42] R. Prix, *The F-statistic and its implementation in ComputeFStatistic v2*, Tech. Rep. (2010) (LIGO-T0900149-v2).

[43] A. Mercer *et al.*, "LSC Algorithm Library - LALSuite," free software (2012).

[44] H. J. Pletsch and B. Allen, Phys. Rev. Lett. **103**, 181102 (2009), arXiv:0906.0023 [gr-qc].

[45] B. Allen *et al.*, "Einstein@Home distributed computing project," .

[46] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **80**, 042003 (2009), arXiv:0905.1705 [gr-qc].